Essays in Industrial Organization under Digital Economy

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Doktor der Wirtschaftswissenschaft)

eingereicht an der

Wirtschaftswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von

Tianchi Li, M. Sc.

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Dr. Daniel Klapper

Gutachter/Gutachterin:
1. Prof. Dr. Roland Strausz
2. Prof. Dr. Anja Schöttner

Eingereicht am: 31. Dezember 2020
Tag des Kolloquiums: 15. Juni 2021

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Roland Strausz for his support and guidance throughout my PhD study. Next, I am grateful to Anja Schöttner for her feedbacks and suggestions as the second supervisor. I also want to show my great thank to Helmut Bester for his always helpful advice at various stages of these three chapters.

Moreover, two coauthors, Friederike Heiny and Michel Tolksdorf, deserve a special thank from me. It is my pleasure to collaborate with you on chapter two and I really appreciate this experience.

Futhermore, I would like to show my thanks to all my colleagues and participants in Berlin Microeconomic Theory Colloquium. Your valuable comments and suggestions make this dissertation possible. In particular, I thank Julian Emmler, Tobias Gamp, Slobodan Sudaric, Colin von Negenborn and Ran Weksler for their insightful discussions and helpful advice.

In addition, a special thank should be given to our administrative team – Regine Hallmann, Viviana Lalli, Kristin Schwier, Myrna Selling and Sandra Uzman. Thanks for your work and help. I also want to express my thanks to all the participants and discussants in different workshops, seminars and conferences. Financial support by DFG (Deutsche Forschungsgemeinschaft) and WWG (Wirtschaftswissenschaftliche Gesellschaft) is gratefully acknowledged.

Finally, I am deeply grateful to my parents for their continued and unconditional support. This dissertation is dedicated to my beloved grandmother, your love leaves memories that live forever in my heart.

Abstract

This thesis analyzes three distinct markets under digital economy and offers insight on the relevant welfare implications and policy suggestions. The first chapter focuses on the crowdfunding market and demonstrates that asymmetric information leads to a higher pre-ordering price which lessens the risk of moral hazard. It also points out the trade-off between making the crowdfunding succeed and preventing the entrepreneur from running away after the success. The second chapter combines theoretical analysis with laboratory experiment to explore the effect of General Data Protection Regulation (GDPR) on consumer's privacy choice. It shows that mandating data sharing among firms leads consumers to give up more privacy and increases consumer welfare, which acts as a good supplement for European Commission's debate on the mandated data sharing in specific sectors. The last chapter discusses the on-demand platform in taxi market where the price is flexible and adjusted by aggregate demand and supply. Contrary to common sense, the result suggests that enough new entrants may not solve the problem of undersupply when demand surges. Social planner, on-demand platform and service providers have different incentives regarding when to enter the market.

Zusammenfassung

Diese Dissertation analysiert drei unterschiedliche Märkte im Rahmen der digitalen Wirtschaft und bietet einen Einblick in die relevanten Wohlfahrtsimplikationen und politischen Vorschläge. Das erste Kapitel konzentriert sich auf den Crowdfunding-Markt und zeigt, dass asymmetrische Informationen zu einem höheren Preis bei Vorbestllung führen, was das Risiko von Moral Hazard verringert. Es weist auch auf den Konflikt zwischen dem Erfolg des Crowdfundings und dem Verhindern der Flucht des Entrepreneurs nach dem Erfolg hin. Das zweite Kapitel kombiniert theoretische Analysen mit einem Laborexperimente, um die Auswirkungen der Datenschutz-Grundverordnung (DSGVO) auf die Wahl von Verbrauchern bezüglich ihrer Privatsphäre zu untersuchen. Es zeigt, dass eine vorgeschriebene gemeinsame Nutzung von Daten zwischen Unternehmen dazu führt, dass Verbraucher vermehrt ihre Daten zur Verfügung stellen und dass Verbraucherwohl steigt. Dies ist eine gute Ergänzung für die Debatte der Europäischen Kommission über eine vorgeschriebene gemeinsame Nutzung von Daten in bestimmten Sektoren darstellt. Im letzten Kapitel wird die On-Demand-Plattform im Taximarkt erörtert, bei der der Preis flexibel ist und an die aggregierte Nachfrage und das Angebot angepasst wird. Gegen die eigene Intuition deutet das Ergebnis darauf hin, dass das Problem von zu wenig Angebot bei steigender Nachfrage möglicherweise nicht durch genügend neue Marktteilnehmer gelöst wird. Sozialplaner, On-Demand-Plattform und Dienstleister haben unterschiedliche Anreize für den Markteintritt.

Contents

List of Figures & Tables

Introduction

Digital economy, also known as Internet Economy or Web Economy, is reshaping our daily life. These economic activities result from billions of everyday online connections among people, businesses, devices, data and processes.[1] With such digital transformations, Facebook and Instagram let people from different areas share their life with others. Uber and Airbnb make travelling much easier and more convenient. Kickstarter and Indiegogo connect great business ideas with monetary support more closely. Alibaba and Amazon solve hundreds of thousands of people's urgent needs during covid-19 pandemic time.

Three independent chapters of this thesis analyze three distinct markets associated with digitalization, including the cases of crowdfunding, behavior-based price discrimination and on-demand platform. In each case, new characteristics brought by digitalization make it different from the traditional market, but also expose some potential problems. For instance, internet-based crowdfunding contains asymmetric information between the entrepreneur and backers, online behavior-based pricing leads consumers to think more cautiously about their privacy choices, and on-demand platform makes it possible to use the flexible price to extract more profits. This thesis tries to understand the economic interactions behind these digital markets and offers the relevant insight on the welfare implications and policy suggestions.

Chapter 1 focuses on the supply side in digital economy, by taking crowdfunding as an example and discussing one problem brought by digitalization – asymmetric information. It shows that how the supplier signals the product quality when asymmetric information exists. This chapter treats a reward-based crowdfunding as a two-stage game, combining the crowdfunding process with the retail market. In the benchmark case without uncertainty, the trade-off between making the crowdfunding succeed and preventing the entrepreneur from running away after the success is pointed out. It also proves that there exists a threshold value of investment requirement, over which

[1]See https://www2.deloitte.com/mt/en/pages/technology/articles/mt-what-is-digital-economy.html (accessed on December 18, 2020).

all the crowdfunding fails. In the case with uncertainty about the product quality, a comparison is made between three scenarios with full information, symmetric uncertainty and asymmetric information. It is demonstrated that under the asymmetric information where the entrepreneur has the extra information about the quality, he raises the preordering price to signal the high quality, which also lessens the risk of moral hazard. Depending on the extra utility of longer usage by pledging, signaling may not influence the success of projects but lowers down the total expected profits.

Chapter 2 concentrates on the customer's behavior, analyzing consumers' privacy choice concerning their data when the privacy is endogenized. With the implementation of EU's General Data Protection Regulation (GDPR), consumers are provided with an option to decide whether to reveal private information in form of cookies. This chapter studies the endogenous decision in a duopoly model with behavior-based pricing and an experiment is conducted to test the theoretical predictions. Contrasting two data environments, a unique pure-strategy pooling equilibrium is derived for each environment. It finds that all consumers share their data in an open data environment while no consumers share data in an exclusive environment. In the experiment, however, some subjects readily share their data in the exclusive treatment. In the open data treatment subjects predominantly act as predicted by the theory. A mandated data sharing policy among firms leads consumers to share more data and increases consumer welfare. This helps policy makers to make an informed decision about data protection policies and acts as a good supplement for the debate on the mandated data sharing in specific sectors.

Different from the first two chapters, Chapter 3 puts emphasis on the intermediary and analyzes a new market associated with digitalization – on-demand platform. Similar to the two-sided market, on-demand platform acts as an intermediary to connect the demand and supply side. This chapter checks an on-demand platform with independent service providers who can freely enter and move between the markets. It considers a two-period model with two markets where the demand randomly surges in one of the markets at the second period. Contrary to common sense, the result shows that even with enough new entrants, the undersupply may still exist. This takes place only when the new entry happens at the first period and the total supply exceeds the total demand within a certain range. Concerning the welfare, new entrants prefer to enter the market in the beginning, while both the social planner and platform have the incentive to postpone the new entry to the period when the demand surges. Limiting the quota, using the bonus to encourage late entry, or applying the surge pricing on the market without surge demand are the possible ways for the platform to solve the problem of undersupply.

Chapter 1

Crowdfunding with Asymmetric Information

Based on Li (2020a).

1.1 Introduction

Crowdfunding is the practice of funding a project or venture by raising small amounts of money from a large number of people, typically via internet. Entrepreneurs, who are lack of money, start their own projects online via the crowdfunding platforms such as Kickstarter or Indiegogo. They post the detailed descriptions of their projects, together with the total money they need and the reward they can offer to those who make the contribution to them. Consumers, who are interested in the project, decide whether to support and how much to pledge. If eventually the amount of money an entrepreneur collects exceeds the target he sets, the crowdfunding succeeds and the entrepreneur gets all the money the consumers pledged to start this project. Otherwise, the crowdfunding fails and all the money would be paid back to those who pledged before.

Other than the traditional funding methods such as business loans and venture capital, crowdfunding provides a novel way for the small investors engaging in supporting the projects. Statistics show that until December 18, 2020, totally over USD 5.4 billion have been raised with 193,900 successfully funded projects on Kickstarter.[1] The market size of crowdfunding has been larger than angel funds and will soon catch up with the venture capital (Chang, 2020). It also provides more functions than the traditional fundings. For instance, the survey by Mollick and Kuppuswamy (2014) shows that most entrepreneurs

[1]Please see https://www.kickstarter.com/help/stats (accessed on December 18, 2020).

initiating the projects via crowdfunding would like to figure out the potential demands for their products. Marketing and connecting to the correct community are ranked as the second and the third reasons. Raising funds is just put in the fourth place. All of these concerns make crowdfunding distinctive from other funding methods.

There are various types of crowdfunding, such as reward-based crowdfunding, equity-based crowdfunding and social crowdfunding. In academic analysis, researchers focus more on reward-based and equity-based crowdfunding, in which the future product or the share of future profits is provided to the consumers as the gratitude for support (Belleflamme et al., 2014). Two mechanisms are commonly applied in crowdfunding, all-or-nothing (AON) and keep-it-all (KIA) mechanisms. The comparisons between these two mechanisms are made by several articles (Chang, 2020, Cumming, Leboeuf and Schwienbacher, 2020) and a general agreement has been reached that AON generates more profit than KIA from both theoretical and empirical sides. Overall, the reward-based crowdfunding with all-or-nothing mechanism is the most common format in practice, which is also the focus in this chapter.

On the other hand, during the implementation of the crowdfunding projects, many problems have been exposed. For instance, Mollick (2014) shows that many unexpected issues such as manufacturing and certification problems, and changes in scale or scope, always occur in the crowdfunding projects, which lead to the delay of shipping in 75% of the cases. Moreover, Belleflamme et al. (2015) point out several problems with respect to asymmetric information, including hidden information and hidden action. This is mainly caused by the fact that unlike the traditional fundings, due diligence cannot be carried out in crowdfunding and not all the information can be disclosed truthfully. Additionally, reported and perceived frauds sometimes take place and form a major concern regarding crowdfunding (Cumming, Hornuf, Karami and Schweizer, 2020).

Concerning these potential problems brought by crowdfunding, several articles have made the relevant analyses from both theoretical and empirical sides, especially about moral hazard and asymmetric information. However, it is lack of the paper jointly discussing these issues. Therefore, we build a theoretical model in this chapter and address the following research questions: what is the effect of moral hazard on the crowdfunding? What changes will it make when the asymmetric information is taken into account? We treat the crowdfunding as a signaling game and want to figure out the entrepreneur's optimal strategy accordingly.

In our model, we focus on the reward-based crowdfunding with two stages. The first stage is the crowdfunding process, where the entrepreneur initiates the project and tries to collect enough pledges to make it feasible. Considering the all-or-nothing

(AON) mechanism that we apply, the entrepreneur receives the money only when the crowdfunding succeeds. In the second stage, the entrepreneur starts mass production and sells the good on the retail market. The benchmark case without uncertainty shows the trade-off between making the crowdfunding succeed in the first stage and preventing the entrepreneur from running away after the success. We demonstrate that in our setup there exists a threshold value of investment requirement, over which all the crowdfunding fails. Then we introduce the uncertainty about product quality and check three different cases: full information, symmetric uncertainty and asymmetric information. We show that the case of symmetric uncertainty, in which both the entrepreneur and consumers do not have extra information, is identical to the benchmark case. Under the asymmetric information where only the entrepreneur knows the quality, the entrepreneur with high-quality product raises the prices to make the signal, which also reduces the risk of moral hazard. Depending on the extra utility of longer usage by pledging, signaling may not influence the success of projects but lowers down the total profits.

1.2 Related Literature

Many recent articles discuss the topics about crowdfunding, combining theoretical modeling with empirical analysis. Belleflamme et al. (2015) summarize various issues that the crowdfunding platform may encounter, including the asymmetric information and fraud. They discuss how consumers screen and how the entrepreneur signals in the crowdfunding, however, without introducing any concrete models to describe the certain scenario. This chapter uses a specific model based on Belleflamme et al. (2014) to consider both the asymmetric information and moral hazard problems, some of which can also be found in Strausz (2017), Schwienbacher (2018), Ellman and Hurkens (2019), Chang (2020), and Chemla and Tinn (2020). It is also related to the topics of price discrimination mentioned by Ellman and Hurkens (2019). Additionally, some empirical papers provide different insight into such problems (Mollick and Kuppuswamy, 2014, Ahlers et al., 2015, Kunz et al., 2017, Cumming, Hornuf, Karami and Schweizer, 2020).

The paper closest to this chapter is Belleflamme et al. (2014). Both of us assume that consumers could get extra utility by pledging, however, with different reasons. They consider from the perspective of community, while our assumption is based on the longer usage of the new products, which is similar to Agrawal et al. (2014). In Belleflamme et al. (2014), the authors make the comparisons between reward-based and equity-based crowdfunding without considering the moral hazard problem, while we only focus on the case of reward-based crowdfunding by taking moral hazard problem into account.

Belleflamme et al. (2014) also briefly mention the asymmetric information, but do not provide a further analysis. They try to identify how the asymmetric information affects the largest project size, in other words, the type of the projects that the entrepreneur could finance. They calculate the different ranges within which either both types can be supported or only one type is able to be operated. Different from their perspective, we treat the crowdfunding as a signaling game, where the entrepreneur can use different prices to signal the quality of their products.

Considering asymmetric information, Belleflamme et al. (2014) firstly make a brief analysis on crowdfunding. Chang (2020) discusses the case in which the consumers can privately receive a signal about the project valuation and then make the relevant decisions. The author focuses on the comparison between fixed and flexible fundings, and demontrates that without financial constraint the entrepreneur could extract full surplus. Hakenes and Schlegel (2014) study the situation where the households get a noisy signal of the firm's quality by incurring a certain cost. They analyze the mixed strategy while we just look at the pure strategies. Kunz et al. (2017) mention more types of signals that could be used in a realistic reward-based crowdfunding and empirically check the effect of different factors on the success rates. In this chapter we discuss the case where only the entrepreneur knows the information about the product and uses different prices to signal the quality. Agrawal et al. (2014), Belleflamme et al. (2015), and Ahlers et al. (2015) propose some other ways in coping with the asymmetric information.

Another issue concerning crowdfunding is moral hazard. Whatever the format is, in principle, the entrepreneur can run away with the money that he has collected from the pledges if the crowdfunding succeeds. This may lead to the severe moral hazard problems. Strausz (2017) proposes that the platform can use the deferred payments to lessen such pressure and should give as little information as possible about the exact number. The author also states that the moral hazard is a first-order problem and private cost information is of second order, which provides us a good reason to attach the importance on this side and incorporate moral hazard into our model. Chemla and Tinn (2020) treat crowdfunding as a learning device to verify the total demand, together with reducing moral hazard problem, especially for the firm facing highly uncertain demand. They also predict that the pre-ordering price should be at par or at discount compared with the retail price. This differs from our result, however, no clear evidence is found by them to show this pattern. Cumming, Hornuf, Karami and Schweizer (2020) point out the behavioral characteristics of moral hazard, together with more types of risks that the crowdfunding project may face, such as misrepresentation and over-optimism. We contribute to this topic by integrating moral hazard problem with asymmetric information, demonstrating that signaling the product quality can reduce the risk of moral hazard.

Other topics, such as demand uncertainty and price discrimination, have also been prevalently discussed in the related literature. Strausz (2017) looks at the problem from mechanism designer's perspective and shows that crowdfunding can solve the problem of demand uncertainty. Schwienbacher (2018), Ellman and Hurkens (2019), and Chang (2020) model the benefit of solving demand uncertainty as well. In Belleflamme et al. (2014) and Ellman and Hurkens (2019) the authors treat crowdfunding as a practice of price discrimination, which support our findings that the entrepreneur sets different pre-ordering and retail prices, in order to maximize his expected profits. Crowdfunding-related topics also attract empirical analysts' attentions. For instance, Cumming, Leboeuf and Schwienbacher (2020) make comparisons between all-or-nothing (AON) and keep-it-all (KIA) on the raising capital and success rate, showing that AON discloses more information. They also show that signaling the product quality is common, especially for the entrepreneur with promising projects, which also backs up our result that signaling may not influence the success of projects if the extra utility is large. Many other papers, such as Mollick and Kuppuswamy (2014), Mollick (2014), and Ahlers et al. (2015) also list out some interesting facts about high success rates, little moral hazard problems, and the common delay of delivery in crowdfunding. Their results act as good complements to the theoretical analyses.

1.3 Model

There are three types of players in the model: a continuum of consumers, an entrepreneur and the crowdfunding platform.

Consumers

There is a continuum of consumers with total mass of one. Their valuation of the product, v, is uniformly distributed between 0 and 1. Each consumer would like to buy one unit of the product in either the first or the second stage. We assume that consumers could get extra utility by pledging, which is proportionate to their valuation. We use α to denote this proportion, meaning that if a consumer's valuation of the product is v, then by pledging his payoff would be $(1+\alpha)v$. This assumption is based on the fact that consumers benefit from the longer usage of the product from crowdfunding. To the contrary, if he decides not to pledge and wait until the retail market, the payoff is just his true valuation of the product, v. Moreover, there is no time preference for consumers, and no discount factor will be taken into consideration in the model.

Entrepreneur

The entrepreneur has a very innovative idea, which ensures that he acts as a monopolist in the market. However, he is lack of money,[2] so he has to rely on the crowdfunding to collect the amount of money he needs. If the crowdfunding succeeds, he can get all the pledges and invest I as a sunk cost before the production. The marginal cost after investing I is constant, which is normalized to zero. When the crowdfunding fails, the entrepreneur gets nothing and the game ends.

Crowdfunding Platform

In our model, platform just offers a place for crowdfunding. It won't charge any fees from consumers or the entrepreneur, and it has nothing to do with the following transactions on the retail market. Therefore, we don't consider any strategic problems regarding the platform in this chapter, *i.e.*, how to set the optimal fees for the platform.

Similar to the biggest online platform Kickstarter,[3] we consider a reward-based all-or-nothing crowdfunding mechanism in this model. "Reward-based" means that if the crowdfunding eventually succeeds, all the consumers who have pledged will get one unit of product as a reward.[4] And "all-or-nothing" shows that when the crowdfunding fails in the end, all the money would be refunded to the consumers who have pledged. On the other hand, if the crowdfunding succeeds, the entrepreneur can get all the pledges.

Timing

We consider a two-stage game. The first stage is the crowdfunding process, in the beginning of which the entrepreneur initiates a reward-based crowdfunding by announcing the investment requirement I and the pledging level p_1. Then consumers come and decide whether to pledge now or wait until the second stage. If the total amount of pledges P exceeds I, the crowdfunding process succeeds and the entrepreneur gets all the pledges P. Then he invests I to start production, where I can be treated as sunk cost. Finally he needs to keep his promise, delivering one unit of the product to all the consumers

[2] In the model we assume that initially the entrepreneur has no money.

[3] More details can be found on the website *www.kickstarter.com*.

[4] In reality, the entrepreneur usually offers different pledging levels for consumers on the crowdfunding platform. Therefore, the rewards for consumers depend on the pledging levels they have chosen, which are generally different. In order to simplify the analysis, we assume that there just exists one type of the pledging level, and the same for reward.

who have pledged.[5] On the other hand, if the investment requirement I has not been satisfied ($P < I$), the game ends and all the money will be paid back to consumers.

Based on the success of crowdfunding process, the game moves to the second stage. In this phase, the entrepreneur starts mass production, sets the unit price p_2, and sells his products on the retail market. Those who didn't pledge in the first stage can buy the product now if their true valuations are not less than the unit price, that is $v \geq p_2$. In our model, we assume that there is no new arrival of consumers in the second stage. In other words, the entrepreneur is facing the residual demands on the retail market.

This model covers not only the traditional crowdfunding process, but the following retail market as well. In this two-stage game, the entrepreneur has to set p_1 and p_2 sequentially, in order to maximize his total expected profits. Meanwhile, he also has two concerns, feasibility constraint and moral hazard problem. Feasibility constraint means that $P \geq I$,[6] which is the necessary and sufficient condition for the success of crowdfunding process. Additionally, we consider moral hazard problem in this model, meaning that the entrepreneur may get all the money from the successful crowdfunding and then run away. In order to aviod such a moral hazard problem, the net profit he can get from the second stage should exceed the money he would obtain by running away. In the following sections, we will analyze all these problems in detail.

1.4 Equilibrium Analysis

In this section, we analyze the model from two aspects. On the one hand, we check the benchmark case without uncertainty. We firstly show the unconstrained optimal solution, and then take both the feasibility and incentive constraints into account, in order to identify the relevant changes. On the other hand, we incorporate the uncertainty about product quality into the model and make the comparisons between three scenarios: full information, symmetric uncertainty and asymmetric information. Regarding the last scenario with asymmetric information, we treat crowdfunding as a signaling game and try to find out the entrepreneur's optimal strategy.

1.4.1 Benchmark

As a benchmark model, we consider a two-stage game without any uncertainty and asymmetric information. In the beginning of the first stage, the entrepreneur announces

[5] Please note that in the benchmark model, we do not consider any cases with asymmetric information or uncertainty. We will make the relevant analyses in the following sections.

[6] This is in accordance with the assumption that initially the entrepreneur has no money.

(p_1, I) for the project. If the crowdfunding succeeds, he sets the retail price p_2 in the second stage. Generally, p_1 is different from p_2 because p_1 can be treated as the pre-ordering price for one unit of product and p_2 is the retail price. In order to solve this problem, we firstly figure out the consumer who is indifferent between pledging in the first stage and waiting until the second stage, and then use the backward induction to find out the equilibrium.

Based on the setup, if a consumer pledges in the first stage, the payoff he would get is $(1+\alpha)v - p_1$. If this consumer decides to wait and buy one unit in the second stage, the payoff will be $v - p_2$. For the consumer who is indifferent between pledging in the first stage and buying in the second stage, the following equivalence should hold

$$(1+\alpha)\bar{v} - p_1 = \bar{v} - p_2 \tag{1.1}$$

where $\bar{v}$ can be expressed as $\bar{v} = \frac{p_1 - p_2}{\alpha}$. For those whose valuations are above $\bar{v}$, they prefer to pledge in the first stage. To the contrary, consumers with the valuations below $\bar{v}$ choose to wait until the retail market. Thus, when we consider the second stage, the entrepreneur sets the retail price p_2 to maximize his profit among the residual demands. His objective function is

$$\max_{p_2} \; p_2(\bar{v} - p_2) \tag{1.2}$$

from which we have that $p_2 = \frac{\bar{v}}{2}$. Finally we come back to the first stage, where the entrepreneur chooses p_1 to maximize his total expected profits in two stages:

$$\max_{p_1} \; p_1(1-\bar{v}) + p_2(\bar{v} - p_2). \tag{1.3}$$

By solving (1.1) and (1.2) we can express $\bar{v}$ and p_2 as a function of p_1, that is, $\bar{v} = \frac{2p_1}{2\alpha+1}$ and $p_2 = \frac{p_1}{2\alpha+1}$. Then plugging $\bar{v}$ and p_2 into (1.3), we get the equilibrium outcome:

$$p_1^* = \frac{(2\alpha+1)^2}{8\alpha+2} \qquad p_2^* = \frac{2\alpha+1}{8\alpha+2} \qquad \bar{v}^* = \frac{2\alpha+1}{4\alpha+1}.$$

This is an unconstrained optimal solution without considering feasibility constraint and moral hazard problem. Now we take these two factors into account, and check what changes they will bring on the equilibrium outcome.

Recall that feasibility constraint means $P \geq I$, where I is the investment the entrepreneur has to make before production. Considering the assumption that the entrepreneur initially has no money and the future production totally relies on what he can get from crowdfunding, the feasibility constraint is equivalent to $\pi_1^* = p_1^*(1-\bar{v}^*) \geq I$, where π_1^* represents the revenue from the first stage. By inserting p_1^*, p_2^* and $\bar{v}^*$ into the inequality, we can get

$$\pi_1^* = \alpha \frac{(2\alpha+1)^2}{(4\alpha+1)^2} \geq I \tag{1.4}$$

where Equation (1.4) represents the feasibility constraint of crowdfunding.

Additionally, we need to consider the moral hazard problem. This is an intrinsic problem in crowdfunding, because after the success of crowdfunding, the entrepreneur may have incentive to run away with all the pledges. Under such circumstance, it's almost impossible for consumers to claim their money back, since the entrepreneur could always insist that he has tried hard but failed, and all the money has been spent during the trial. Due to the lack of supervision and punishment mechanism, moral hazard problem cannot be ignored when crowdfunding is considered. Therefore, we need to give the entrepreneur proper incentives, in order to make sure that he won't run away with all the pledges when crowdfunding succeeds. In other words, the entrepreneur could get more from staying in the retail market than from running away, that is, $P + \pi_2^* - I \geq P$. The left hand side of this inequality represents the total benefit for the entrepreneur if he decides to stay in the retail market, where π_2^* is the revenue from the second stage and I is the investment cost. And the right hand side P stands for the benefit he could get by running away.[7] Thus, we can derive the incentive constraint, to make sure that there is no moral hazard problem in crowdfunding. That is, $\pi_2^* \geq I$, or by plugging p_1^*, p_2^* and $\bar{v}^*$ into it, we have

$$\pi_2^* = \frac{1}{4} \frac{(2\alpha+1)^2}{(4\alpha+1)^2} \geq I. \tag{1.5}$$

Equation (1.5) can be treated as the incentive constraint of crowdfunding. Now we need to add both the feasibility and incentive constraints into the optimization problem in Equation (1.3), to find out the constrained equilibrium. Three different scenarios are discussed as follows:

Scenario I $\alpha < \frac{1}{4}$ $(\pi_1^* < \pi_2^*)$

Based on the unconstrained optimal outcome, we have $\pi_1^* < \pi_2^*$ under such circumstance. When the investment requirement I is small, π_1^* and π_2^* are larger than I. Both Equation (1.4) and (1.5) are satisfied, therefore (p_1^*, p_2^*) constitutes an equilibrium. However, with the increase of I, the investment requirement will exceed π_1^*, meaning that the crowdfunding is not feasible. In order to make it succeed, the entrepreneur needs to

[7] Here we assume that by running away, the entrepreneur could get all the pledges. In Strausz (2017), the autor uses αI on the right hand side instead, which means that the entrepreneur can only get part of the pledges if he runs away. We don't consider this situation in this chapter.

increase π_1 by changing his pricing strategy. We can easily check that reducing p_1 leads to the increase of π_1 and the decrease of π_2. And the total profits will also decrease. As an entrepreneur, he will make the feasibility constraint binding, that is $\pi_1 = p_1(1-\bar{v}) = I$, to obtain the maximal profits. Inserting $\bar{v}$ into the binding constraint, we get

$$p_1(1 - \frac{2p_1}{2\alpha+1}) = I. \tag{1.6}$$

Since Equation (1.6) is a quadratic form of p_1, in order to make sure the existence of roots in p_1, we have $(2\alpha+1)^2 - 8(2\alpha+1)I \geq 0$, which is equivalent to $I \leq \frac{1}{8}(2\alpha+1)$. When I is within this interval, we can derive that $p_1 = \frac{2\alpha+1}{4} + \frac{1}{4}\sqrt{(2\alpha+1)(2\alpha+1-8I)}$, which is smaller than p_1^*.

By reducing p_1 to the level mentioned above, the entrepreneur ensures that $\pi_1 = I$ and $\pi_2 \geq I$. However, with the increase of I, the entrepreneur needs to set p_1 lower to satisfy the feasibility constraint, which reduces π_2 as well. There is a threshold value of I, above which the feasibility constraint and incentive constraint cannot be satisfied simultaneously. Considering this threshold, $\pi_1 = \pi_2 = I$ should hold, which means that $p_1(1-\frac{2p_1}{2\alpha+1}) = I = (\frac{p_1}{2\alpha+1})^2$. We can calculate that $p_1 = \frac{(2\alpha+1)^2}{4\alpha+3}$, and also check that under such circumstance $I = (\frac{2\alpha+1}{4\alpha+3})^2 < \frac{1}{8}(2\alpha+1)$, which satisfies the prerequisite condition from Equation (1.6). When I is larger than this threshold value, entrepreneur will prefer to run away rather than stay in the retail market.

In reality, if the investment requirement of the crowdfunding is too high, the entrepreneur needs to lower the pledging level in order to attract more consumers to pledge. However, with more pledges from the first stage and less profit in the second stage, the moral hazard problem may occur. When I exceeds a threshold, either Equation (1.4) or (1.5) will be violated, therefore the crowdfunding fails.

Scenario II $\alpha > \frac{1}{4}$ $(\pi_1^* > \pi_2^*)$

In this scenario we have $\pi_1^* > \pi_2^*$. Similarly, when I is small, both π_1^* and π_2^* are larger than I, satisfying Equation (1.4) and (1.5). (p_1^*, p_2^*) remains as an equilibrium. Opposite to Scenario I, with the increase of I, it would be that $\pi_2^* < I$, showing the existence of moral hazard problem. Thus, the entrepreneur needs to increase π_2 to solve such a problem. It is easy to see that he will increase p_1 to make the incentive constraint binding, that is $\pi_2 = p_2(\bar{v} - p_2) = I$, in order to gain the maximal profits. So we have that

$$(\frac{p_1}{2\alpha+1})^2 = I. \tag{1.7}$$

By solving Equation (1.7), we get that $p_1 = (2\alpha+1)\sqrt{I}$, which is larger than p_1^*. Therefore, the entrepreneur increases p_1 to $(2\alpha+1)\sqrt{I}$, making $\pi_1 \geq I$ and $\pi_2 = I$, in order to eliminate the moral hazard problem. Similarly, there is a threshold value of I, over which the feasibility and incentive constraints cannot hold simultaneously. Let $\pi_1 = \pi_2 = I$, that is, $p_1(1-\frac{2p_1}{2\alpha+1}) = I = (\frac{p_1}{2\alpha+1})^2$. We can calculate that this threshold is the same as Scenario I. In other words, the threshold is $I = (\frac{2\alpha+1}{4\alpha+3})^2$ with the price $p_1 = \frac{(2\alpha+1)^2}{4\alpha+3}$.

To sum up, in this scenario, with the increase of I, the entrepreneur has to raise p_1, in order to eliminate the potential moral hazard problem. Concerning crowdfunding, we can describe this strategy as lowering the profit in the first stage. By doing so, the total amount of money the entrepreneur could get from crowdfunding is reduced, and therefore consumers won't worry about the moral hazard problem. Meanwhile, the profit in the second stage increases, which incentivizes the entrepreneur to stay in the retail market, not running away. However, when I is larger than the threshold value, either the feasibility constraint or the incentive constraint is violated, so the game ends and crowdfunding fails.

Scenario III $\alpha = \frac{1}{4}$ $(\pi_1^* = \pi_2^*)$

Lastly we check the scenario where $\alpha = \frac{1}{4}$. In this situation, $\pi_1^* = \pi_2^*$. When I is small, both the feasibility and incentive constraints are satisfied by the unconstrained optimal results and the entrepreneur has no incentive to deviate. Thus, (p_1^*, p_2^*) forms an equilibrium. When I exceeds π_1^* (π_2^*), it's impossible to increase both π_1^* and π_2^* simultaneously, since the unconstrained solution maximizes the sum of $\pi_1^* + \pi_2^*$.[8] It means that at least one of the constraints cannot be satisfied, no matter how the entrepreneur adjusts his strategy. If the feasibility constraint is violated, crowdfunding will fail eventually. Otherwise, if the incentive constraint is not met, consumers won't pledge because they expect that the entrepreneur would run away with all the pledges. Thus, there is no equilibrium when $I > \pi_1^*$ (π_2^*).

Proposition 1.1. *When the investment requirement I is low, $(p_1^*, p_2^*) = (\frac{(2\alpha+1)^2}{8\alpha+2}, \frac{2\alpha+1}{8\alpha+2})$ constitutes the equilibrium in this game. However, with the increase of I, entrepreneur has to reduce p_1 if $\alpha < \frac{1}{4}$, or increase p_1 if $\alpha > \frac{1}{4}$, in order to satisfy both the feasibility and incentive constraints, leading to the decrease of the total profits. There exists a threshold value of I, over which all the crowdfunding will fail, no matter how large α is.*

Proposition 1.1 summarizes the results in the benchmark model, where neither uncertainty nor asymmetric information is considered. With the different levels of investment requirement, there exists a trade-off between making the crowdfunding succeed in the

[8]Strictly speaking, the equilibrium results maximize the sum of $\pi_1^* + \pi_2^* - I$.

first stage and preventing the entrepreneur from running away after the success. Increasing the profit in the first stage ensures the success of crowdfunding, but also leads to the potential moral hazard problem. The larger π_1 is, the higher incentive the entrepreneur would have to run away. To the contrary, more profit in the second stage results in the lack of pledges from the first stage. The larger π_2 is, the less feasible the crowdfunding will be. This trade-off generally lowers the total profits and causes the failure of crowdfunding when the investment cost I exceeds a threshold value.[9]

1.4.2 Uncertainty and Asymmetric Information

In this section, we take uncertainty and asymmetric information into account and analyze how these factors affect the crowdfunding. Regarding a crowdfunding project, consumers have to decide whether to pledge or wait until the second stage. Their uncertainties lie in two aspects: price and quality. They don't know the future price, and also wonder whether the product will have good quality or not. Since price is the control variable in our model, we focus on the uncertainty about quality.[10] Throughout the following analysis, we assume that there are two types of products: high-quality and low-quality. Consumer's valuation is $(1 + \Delta)v$ for the high-quality product and $(1 - \Delta)v$ for the low-quality product.

Considering crowdfunding itself, in the very beginning some entrepreneurs don't know how good their final products will be. Although they have great ideas, crowdfunding seems to be a trial for them. Therefore, neither the entrepreneur nor consumers know whether it will eventually be a high-quality product or not. We call it *symmetric uncertainty* case, in which no one is sure about quality. To the contrary, some projects have obtained success and those entrepreneurs also have accumulated their own reputation before the crowdfunding. Thus, both the entrepreneur and consumers are familiar with such products and the aim of crowdfunding is to get the financial support for the mass production. We define this as *full information* case, where everyone knows the product quality in the beginning of crowdfunding. Moreover, in some cases, entrepreneurs have more information than consumers. For instance, at Kickstarter, we can find the detailed descriptions about certain products as well as the related introductory videos, all of which show that entrepreneurs should know more than consumers. Therefore, asymmetry exists in such a crowdfunding, which we call *asymmetric information* case.

[9] We provide two examples in Appendix to show more details about the benchmark case.

[10] The analysis about the commitment to the prices is left in Discussion and Appendix.

1.4.2.1 Full Information

We start from the case of full information, where everyone knows the exact product quality in the beginning of crowdfunding. If it is a high-quality product, consumer's valuation will be $(1+\Delta)v$ and the entrepreneur sets p_2^h at $t=2$ to maximize $p_2^h(\bar{v^h}-\frac{p_2^h}{1+\Delta})$, where $\bar{v^h}$ represents the valuation of the consumer who is indifferent between pledging in the first stage and waiting until the second stage. So, we have that $p_2^h=\frac{1+\Delta}{2}\bar{v}$. On the other hand, if the quality of the product is low, consumer's valuation will be $(1-\Delta)v$ and the entrepreneur's objective function becomes $p_2^l(\bar{v^l}-\frac{p_2^l}{1-\Delta})$ in the second stage, where we can derive $p_2^l=\frac{1-\Delta}{2}\bar{v^l}$. Then we move back to the first stage to figure out the condition for the indifferent consumer. For a high-quality product, if a consumer pledges his payoff would be $(1+\Delta)(1+\alpha)\bar{v^h}-p_1^h$. Otherwise, if he waits until the second stage, he could get $(1+\Delta)\bar{v^h}-p_2^h$. By equalizing these two payoffs, that is,

$$(1+\Delta)(1+\alpha)\bar{v^h}-p_1^h=(1+\Delta)\bar{v^h}-p_2^h, \tag{1.8}$$

we get that $\bar{v^h}=\frac{2p_1^h}{(1+\Delta)(2\alpha+1)}$. Inserting this into p_2^h, we have that $p_2^h=\frac{p_1^h}{2\alpha+1}$. Finally, we consider the entrepreneur's objective function. Since this is a full information case, he only needs to maximize the total expected profits in two stages, that is,

$$\max_{p_1^h}\ p_1^h(1-\bar{v^h})+p_2^h(\bar{v^h}-\frac{p_2^h}{1+\Delta}). \tag{1.9}$$

By plugging p_2^h and $\bar{v^h}$ into the optimization problem, we could easily solve that $p_1^h=\frac{(1+\Delta)(2\alpha+1)^2}{8\alpha+2}$, $p_2^h=\frac{(1+\Delta)(2\alpha+1)}{8\alpha+2}$ and $\bar{v^h}=\frac{2\alpha+1}{4\alpha+1}$. In the end, we calculate that $\pi_1^h=\alpha(1+\Delta)\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$, $\pi_2^h=\frac{1+\Delta}{4}\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$ and $\pi_h=\pi_1^h+\pi_2^h=\frac{1+\Delta}{4}\frac{(2\alpha+1)^2}{4\alpha+1}$.

Following the same way, we check the scenario with low-quality product.[11] We can derive that $p_1^l=\frac{(1-\Delta)(2\alpha+1)^2}{8\alpha+2}$, $p_2^l=\frac{(1-\Delta)(2\alpha+1)}{8\alpha+2}$ and $\bar{v^l}=\frac{2\alpha+1}{4\alpha+1}$, where $\bar{v^l}$ stands for the indifferent consumer in the first stage. Moreover, $\pi_1^l=\alpha(1-\Delta)\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$, $\pi_2^l=\frac{1-\Delta}{4}\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$ and $\pi_l=\pi_1^l+\pi_2^l=\frac{1-\Delta}{4}\frac{(2\alpha+1)^2}{4\alpha+1}$.

1.4.2.2 Symmetric Uncertainty

Now let's look at the case with symmetric uncertainty. As introduced above, symmetric uncertainty means that both the entrepreneur and consumers don't know the quality of the final product. Therefore, we assume that there is a probability of $\frac{1}{2}$ that the final product has high quality and consumer's valuation is $(1+\Delta)v$. On the other hand,

[11] The details of the analysis are shown in Appendix.

there is another probability of $\frac{1}{2}$ that the product proves to be with low quality, where consumer's valuation is $(1-\Delta)v$. All the information about quality will be revealed after the first stage when the trial has been made. Thus, the timeline is as follows:

t=1: The entrepreneur initiates the crowdfunding project without extra information about quality and announces (p_1, I). Then consumers come and decide whether to pledge or wait. If crowdfunding succeeds, the entrepreneur receives all the pledges and invests I to start production. After the trial, the quality of final product is revealed. Otherwise, if crowdfunding fails, all the pledges will be paid back to consumers and the game ends.

t=2: After the success of crowdfunding, everyone knows the quality of product. The entrepreneur sets unit price p_2 and the remaining consumers decide whether to buy on the retail market.

In order to solve this two-stage game, we use the backward induction again. At $t = 2$, it is a full information case. As shown before, the entrepreneur chooses p_2^h to maximize $p_2^h(\bar{v^h} - \frac{p_2^h}{1+\Delta})$ where $\bar{v^h}$ represents the marginal consumer in the first stage, and we have that $p_2^h = \frac{1+\Delta}{2}\bar{v^h}$. Similarly we can derive $p_2^l = \frac{1-\Delta}{2}\bar{v^l}$.

When $t = 1$, we need to firstly find out the consumer who is indifferent between these two stages. By pledging in the first stage, the expected payoff will be

$$\frac{1}{2}(1+\alpha)(1+\Delta)\bar{v} + \frac{1}{2}(1+\alpha)(1-\Delta)\bar{v} - p_1.$$

If this consumer decides to wait and buy in the second stage, his expected payoff will be

$$\frac{1}{2}[(1+\Delta)\bar{v} - p_2^h] + \frac{1}{2}[(1-\Delta)\bar{v} - p_2^l].$$

Therefore, the valuation of the consumer who is indifferent between pledging and waiting satisfies

$$\frac{1}{2}(1+\alpha)(1+\Delta)\bar{v} + \frac{1}{2}(1+\alpha)(1-\Delta)\bar{v} - p_1 = \frac{1}{2}[(1+\Delta)\bar{v} - p_2^h] + \frac{1}{2}[(1-\Delta)\bar{v} - p_2^l]. \quad (1.10)$$

By plugging p_2^h and p_2^l into Equation(1.10), we get $\bar{v} = \frac{2p_1}{2\alpha+1}$, which is identical to the result derived from Equation (1.1) and (1.2) in the benchmark case. Now we look at the entrepreneur's problem, that is, to maximize his expected profits in two stages:

$$\max_{p_1} \; p_1(1-\bar{v}) + \frac{1}{2}p_2^h(\bar{v} - \frac{p_2^h}{1+\Delta}) + \frac{1}{2}p_2^l(\bar{v} - \frac{p_2^l}{1-\Delta}) \quad (1.11)$$

where all of $\bar{v}$, p_2^h and p_2^l can all be expressed as the functions of p_1. By rearranging this equation, we can easily see that the objective function is completely identical to the benchmark model, and the same for the results:

$$p_1^s = p_1^* = \frac{(2\alpha+1)^2}{8\alpha+2} \qquad \bar{v}^s = \bar{v}^* = \frac{2\alpha+1}{4\alpha+1}$$

where p_1^s and $\bar{v}^s$ denote the equilibrium outcomes in symmetric information case. We can further check the values of p_2^h and p_2^l, and verify that π_1^s and $E(\pi_2^s)$ are the same as π_1^* and π_2^* in the benchmark model without any uncertainty. Moreover, same as Equation (1.4) and (1.5), the feasibility constraint and incentive constraint will not change in this case, and therefore the equilibrium result in symmetric uncertainty case keeps the same as the benchmark model.[12]

Proposition 1.2. *The symmetric uncertainty about product quality between the entrepreneur and consumers do not affect the crowdfunding. The equilibrium outcome is identical to the benchmark case.*

Please note that if we just consider the result, the case of symmetric uncertainty is identical to the case where only the entrepreneur knows the information but cannot signal the product quality. Concerning crowdfunding, however, the entrepreneurs with different products should be able to use prices to signal the quality. We discuss this issue in the following section.

1.4.2.3 Asymmetric Information

Asymmetric information is prevalent in crowdfunding. From the previous discussion we can see that some entrepreneurs more or less have more information about the product than consumers. For instance, they know whether their products will eventually have high or low quality, but say nothing about it in the beginning. Thus, we modify the model accordingly in this section. We assume that consumers have the prior belief about the quality as before, that is, there is a probability of $\frac{1}{2}$ that the final product has high quality with consumer's valuation of $(1+\Delta)v$, and probability of $\frac{1}{2}$ with low quality where consumer's valuation is $(1-\Delta)v$. The timeline changes as follows:

[12] We also check the more general case in which there is probability β that the final product has high quality and probability $1-\beta$ of low quality. If $\beta < \frac{1}{2}$, both π_1^s and $E(\pi_2^s)$ will decrease compared with π_1^* and π_2^*. With a more risky project, less consumers make the advanced purchases in the first stage, and the high probability of low-quality product also reduces the expected profit on the retail market. To the contrary, if $\beta > \frac{1}{2}$, both π_1^s and $E(\pi_2^s)$ will increase due to the less risk in the crowdfunding project. More details can be found in Appendix.

t=1: Consumers come to the platform with the identical prior belief. The entrepreneur with the information about quality initiates the crowdfunding project and sets (p_1, I). Consumers update their beliefs based on the observation of the price and decide whether to pledge or wait. If crowdfunding succeeds, the entrepreneur receives all the pledges and invests I to start production. Otherwise, all the pledges will be paid back to consumers and the game ends.

t=2: After the crowdfunding process, consumers know the quality of product. Entrepreneur sets unit price p_2 and the remaining consumers decide whether to buy on the retail market.

Due to the asymmetric information, consumers do not know the product quality in the first stage, and therefore the entrepreneur with low-quality product may pretend to have high quality, in order to obtain more profits. First, we check the scenario in which the entrepreneur with low-quality product mimics high-quality product by using the same pricing strategy as the full information case, to find out whether he could benefit from doing so.

In the first stage, no consumer can tell whether the product has high or low quality. If the entrepreneur with low-quality product sets the pre-ordering price at p_1^h, he will face the indifferent consumer $\bar{v}^h$, where $\bar{v}^h$ comes from the previous case with full information. However, when it comes to the second stage, the quality of product reveals and everyone knows that it is a low-quality product. Thus, the entrepreneur will set the retail price p_2^l to maximize his profit in the second stage, that is,

$$\max_{p_2^l} \; p_2^l(\bar{v}^h - \frac{p_2^l}{1-\Delta}). \tag{1.12}$$

From the previous calculation, we have that $\bar{v}^h = \bar{v}^l$ under the full information. So the entrepreneur's expected profit in the second stage does not change despite the different pricing strategy in the first stage.[13] If the entrepreneur does so, his total profits would be $p_1^h(1-\bar{v}^h) + p_2^l(\bar{v}^h - \frac{p_2^l}{1-\Delta})$, which are higher than what he can get if reporting truthfully. The following corollary summarizes such a deviation:

Corollary 1.1. *With the asymmetric information of product quality, if the entrepreneur offers the low-quality product, he always has the incentive to not tell the truth by setting the price at the high-quality level in the first stage.*

Proof. See Appendix. □

[13]It means that Equation (1.12) is identical to Equation (1.13) in Appendix.

Corollary 1.1 points out the problem with asymmetric information. Consumers will modify their beliefs even though the entrepreneur sets p_1^h, which definitely harms the high-quality entrepreneur. As a response, the entrepreneur with high-quality product could use the different pricing strategy to signal the quality of his product. Therefore, we treat crowdfunding as a signaling game, meanwhile taking the feasibility and incentive constraints into consideration. In the following section we elaborate on this problem and solve it in pure strategies. We focus on the separating equilibrium, in order to find out how to signal in the crowdfunding. The analysis of the pooling case is left in Appendix.

Separating Equilibrium

We have shown that with the asymmetric information the entrepreneur with high-quality product cannot keep the pre-ordering price at p_1^h, since the entrepreneur with low-quality product always has incentive to mimic his strategy. In order to distinguish himself, the entrepreneur with high-quality product has to change his pricing strategy, making sure that no one is willing to mimic. We denote the optimal pricing strategy in separating equilibrium as (p_{1s}^h, p_{2s}^h) and (p_{1s}^l, p_{2s}^l) respectively for high-quality and low-quality entrepreneur. It is straightforward to see that the low-quality entrepreneur will follow the full information case to set $p_{1s}^l = p_1^l$ and $p_{2s}^l = p_2^l$, since it has been a profit-maximal strategy. Considering (p_{1s}^h, p_{2s}^h), it should ensure that the low-quality entrepreneur won't deviate to this strategy and maximize total profits for the high-quality entrepreneur as well. Lemma 1.1 describes such a solution:

Lemma 1.1. *In the optimal separating equilibrium, the entrepreneur with high-quality product raises p_{1s}^h and p_{2s}^h to signal the quality, $p_{1s}^h = \frac{(1+\Delta)^2(2\alpha+1)^2}{2(1+4\alpha+3\Delta+4\alpha\Delta)} + \frac{(1+\Delta)(2\alpha+1)^2}{1+4\alpha+3\Delta+4\alpha\Delta} \cdot \sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}$ and $p_{2s}^h = \frac{(1+\Delta)(1-\Delta)(2\alpha+1)}{2(1+4\alpha+3\Delta+4\alpha\Delta)} + \frac{(1-\Delta)(2\alpha+1)}{1+4\alpha+3\Delta+4\alpha\Delta}\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}$. On the other hand, the low-quality entrepreneur keeps (p_{1s}^l, p_{2s}^l) the same as the full information case.*

Proof. See Appendix. □

Compared with the equilibrium outcome under full information, only the entrepreneur with high-quality product makes the change. Since the high-quality entrepreneur deviates from the optimal strategy and uses the higher prices to signal the quality, he becomes worse off. The following lemma describes such changes on the high-quality entrepreneur brought by the asymmetric information.

Lemma 1.2. *In the optimal separating equilibrium, the profit of the high-quality entrepreneur in the first stage is reduced compared with the case of full information, but* *increased in the second stage. Signaling lessens the total profits, as well as the risk of moral hazard.*

Proof. See Appendix. □

When we check the profits separately between two stages, Lemma 1.2 shows that for the high-quality entrepreneur, the profit from the first stage in the separating equilibrium is less than that in the full information case. To the contrary, the profit from the second stage is larger. The economic implication is, the increase of p_{1s}^h and p_{2s}^h reduces the profit in the crowdfunding process, while raises the profit on the retail market. Considering the fact that the quality reveals after the first stage, the entrepreneur uses this strategy to signal his high quality because he has to rely more on the profit from the second stage under such circumstance, where a low-quality entrepreneur will not do so. More profit in the second stage also reduces the risk of moral hazard, which is, on the other hand, good for the consumers.

In the full information case, $\pi_1^h = \alpha(1+\Delta)\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$ and $\pi_2^h = \frac{1+\Delta}{4}\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$. If $\alpha \leq \frac{1}{4}$, we have $\pi_1^h \leq \pi_2^h$. Lemma 1.2 tells that in the separating equilibrium, high-quality entrepreneur lowers π_1^h in order to signal his quality. Therefore in this case, profit from the first stage will be strictly smaller than that in the second stage. When the feasibility and incentive contraints are taken into account, we can follow the same method as Scenario I in the benchmark model. If the investment requirement I is small, (p_{1s}^h, p_{2s}^h) and (p_{1s}^l, p_{2s}^l) constitute the optimal separating equilibrium. To the contrary, if the investment requirement exceeds π_{1s}^h, high-quality entrepreneur has to reduce p_{1s}^h to ensure the feasibility of crowdfunding. However, it is impossible in the separating equilibrium because signaling requires even lower profit in the first stage compared with the full information case. Thus, high-quality entrepreneur is unable to signal the projects with large investment requirement under the asymmetric information. Additionally, the total profits are reduced as the signaling takes place.

On the other hand, if $\alpha > \frac{1}{4}$, in the optimal separating equilibrium the profit from the first stage may be larger than that from the second stage. For instance, when Δ converges to 0.[14] Taking two constraints into consideration, it is similar to Scenario II in the benchmark model. We have that when I does not exceed the second-stage profit π_{2s}^h, (p_{1s}^h, p_{2s}^h) and (p_{1s}^l, p_{2s}^l) hold as the optimal separating equilibrium. When I is larger than π_{2s}^h, high-quality entrepreneur can increase p_{1s}^h to eliminate the potential

[14] When Δ converges to 0, the results under full information imply that the profit in the first stage is higher. In Appendix we show that π_{1s}^h can be either larger or smaller than π_{2s}^h, depending on the values of α and Δ.

moral hazard problem while still signaling his product quality.[15] Therefore, under such circumstance, the success of high-quality entrepreneur's projects may not be influenced by asymmetric information. He is able to signal the product quality even with high investment requirement. However, the total profits are still driven down under the condition that I is small and no moral hazard problem needs to be considered.

Proposition 1.3. *For the entrepreneur with high-quality product, when the extra utility of longer usage that consumers could get from pledging is small ($\alpha \leq \frac{1}{4}$), signaling the product quality with large investment requirement is prohibited by the asymmetric information inside crowdfunding, and the total profits are driven down. On the other hand, when the extra utility is large ($\alpha > \frac{1}{4}$), asymmetric information may not influence the success of projects, but still lowers the total profits. To the contrary, asymmetric information has no influence on the entrepreneur with low-quality product in the separating equilibrium.*

Proposition 1.3 summarizes that when α is small, asymmetric information prevents high-quality entrepreneur from signaling his project with large investment requirement. But it is not the case when α is larger. This results from the fact that α stands for the extra utility consumers could get by pledging. The larger α is, the more consumers are willing to pledge in the first stage. Asymmetric information leads to a higher risk of supporting the high-quality product, but the extra utility by pledging can mitigate at least part of such influence. Therefore, when α is large enough, asymmetric information won't affect the feasibility of crowdfunding for high-quality entrepreneur. Nevertheless, as a consequence of signaling, the profits are driven down compared with the case under full information.

1.4.2.4 Welfare

Lastly, we compare the welfare in separating equilibrium with the case of full information. No matter how large α is, consumers who purchase the high-quality products become worse off, since they have to pay more for each unit of the good in both stages. Moreover, with the increase of prices, the total number of consumers who buy the high-quality goods decreases,[16] which leads to a reduction in consumer surplus. Concerning the entrepreneur, when I is low, the total profits are reduced due to the signaling of quality. When I is high, depending on the value of α, it will be either infeasible to signal or no change of the

[15] Please note that it is only possible when π_{1s}^h is larger than π_{2s}^h in the optimal separating equilibrium.

[16] This comes directly from the fact that the total profits of high-quality entrepreneur decreases but both pre-ordering and retail prices increase.

profits compared with the full information case. Overall, asymmetric information makes both the entrepreneur and consumers worse off in the case of high-quality products, and has no effect on low-quality products. Information rents make the crowdfunding costly.

For the entrepreneur, building a “correct” community is important. As we have shown, when α is small, asymmetric information will prevent the entrepreneur from signaling the high-quality product with large investment requirement. However, this may not be the case for the higher α. If we treat α as the degree of need, those who really need a certain product should benefit more from the longer usage and therefore have a higher α. Thus, the entrepreneur can try to attract those who really need his product and build a community with higher α. On the other hand, for the goods that need a long waiting time for the production, crowdfunding seems to be a good choice. Consumers can have a longer early trial of those goods, which results in a higher α. Under such circumstance, the entrepreneur can signal the product quality without hurting himself.

Regarding the platform, introducing more mature projects is important. In our model, the value of Δ represents the degree of uncertainty. The more mature the project is, the less uncertain it will be. We have seen that when Δ converges to zero, the inefficiency arises from asymmetric information will be eliminated. Thus, those mature projects can help reduce the uncertainty, and the inefficiency caused by the asymmetric information as well. The evidence shows that among the successful crowdfunding projects, many entrepreneurs have proven their products before. For these entrepreneurs, the aim of crowdfunding is to raise the funds. Other methods, such as pledge cap and deferred payments, should also work. They can efficiently reduce the entrepreneur’s money colletcion in the crowdfunding process and increase the expected profit on retail market.

For those projects that are in the early stage and face a lot of uncertainty, the inefficiency caused by the asymmetric information is large and therefore the success rate would be low. However, as shown from the survey by Mollick and Kuppuswamy (2014), financing the project is only put on the fourth rank considering the purpose of crowdfunding. Checking the potential demands and marketing are treated as the main purposes. From this perspective, uncertainty and asymmetric information do not cause severe troubles on crowdfunding, and as a return, the performance in crowdfunding to some extent helps solve such a problem.

1.5 Conclusion and Discussion

In this chapter, we analyze the entrepreneur’s pricing strategy in a reward-based crowdfunding with all-or-nothing mechanism. In the benchmark case without uncertainty, the optimal pricing strategy should balance the success in the crowdfunding and the avoidance of moral hazard problem afterward. To the contrary, when the uncertainty and asymmetric information are taken into account, the pricing strategy acts as a signal to indicate the product quality. Under the circumstance where only the entrepreneur knows the information about quality, the pre-ordering price is raised by the high-quality entrepreneur to make the signal. This behavior results in the lower profit in the first stage but higher profit in the second stage compared with the case under full information, which on the other hand, lessens the risk of moral hazard. We also demonstrate that depending on the value of the extra utility by pledging, signaling may not affect the feasibility of projects but drive down the total profits.

Throughout the analysis we focus on the uncertainty about product quality, where the entrepreneur sets the pre-ordering price and retail price sequentially. On the other hand, the uncertainty in price is also what the consumers concern. If the retail price is much lower than the pre-ordering price, consumers will prefer to wait until the second stage. Otherwise, more are willing to pledge in the crowdfunding process. As a response, the entrepreneur could make the commitment on the retail price, to mitigate this uncertainty. If we take the benchmark model as an example, we can easily show that without any constraints on feasibility and moral hazard, the entrepreneur would commit to the prices such that all the consumers buy the product in the first stage.[17] This increases the total profits, but lowers both the consumer surplus and social welfare. Even if we take these two constraints into account, we could also demonstrate that committing to the prices makes the entrepreneur strictly better off. However, we should note that the lack of supervision and punishment mechanism makes it hard for consumers to believe such "commitment" in reality.[18] Although we predict a similar result in the case with uncertainty and asymmetric information, the relavant analysis is left for the future research considering that this is not the focal point in this chapter.

We also would like to point out some limitations in our analysis, which may be useful for the future work. For instance, the total demands in the first and the second stages are fixed in our model. It would be interesting to relax this assumption, to verify how the crowdfunding solves the demand uncertainty if both the moral hazard and asymmetric information are considered.

[17]Details about the analysis can be found in Appendix.

[18]For instance, some entrepreneurs use the words such as "you get a certain discount by pledging" to express their commitments. However, we make a simple investigation into these projects and figure out that such commitments do not always work. Among 14 cases with "commitments", only 8 entrepreneurs keep their words. 4 set the earliest retail price higher than the committed price and another 2 lower than the commitment.

Concerning the information discloure, we assume that the product quality can be perfectly revealed after the first stage. In reality, however, imperfect information revelation may take place. In other words, consumers' beliefs about the quality should be updated based on some observable evidence such as updates on the platform, accumulated pledges or the online reviews. Block et al. (2018) have empirically shown that posting the updates regularly can be helpful in raising the money. It might deserve a further study to build such a link between theoretical analysis and empirical observations.

Lastly, throughout the analysis we assume that the entrepreneur truthfully reports the investment requirement I. Strausz (2017) also points out that the entrepreneur's private cost information does not matter if there exists no moral hazard problem, since misreporting brings no benefit to the entrepreneur. However, if we treat crowdfunding as a signaling game, the entrepreneur could in principle also use the investment requirement to signal the product quality, except for the pre-ordering price. Schwienbacher (2018) finds that the entrepreneur may raise more money than necessary when the product is easy to be replicated, so it would also be interesting to verify how to signal the quality by using both the pre-ordering price and reported investment requirement as the tool.

1.6 Appendix

Omitted analysis on benchmark

In this part, we use two graphs ($\alpha = 0.1$ and $\alpha = 0.8$) as the example to show more details about the benchmark case.

If $\alpha < \frac{1}{4}$, as analyzed before we have that $\pi_1^* < \pi_2^*$. When $I \leq \pi_1^* = \alpha\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$, the investment is small and both π_1^* and π_2^* are large enough to satisfy the feasibility and incentive constraints. Thus, the revenue is equivalent to $\pi_1^* + \pi_2^* - I = \frac{(2\alpha+1)^2}{4(4\alpha+1)} - I$, which is linear in I. When $\alpha\frac{(2\alpha+1)^2}{(4\alpha+1)^2} < I \leq (\frac{2\alpha+1}{4\alpha+3})^2$, where $(\frac{2\alpha+1}{4\alpha+3})^2$ is the threshold value of I that we have derived, the entrepreneur should lower p_1 to make the feasibility constraint binding. Under such circumstance, $p_1 = \frac{2\alpha+1}{4} + \frac{1}{4}\sqrt{(2\alpha+1)(2\alpha+1-8I)}$ and the revenue is equal to $(\frac{p_1}{2\alpha+1})^2 = (\frac{1}{4} + \frac{1}{4}\sqrt{\frac{2\alpha+1-8I}{2\alpha+1}})^2$, which is not linear in I. When I exceeds $(\frac{2\alpha+1}{4\alpha+3})^2$, either feasibility constraint or incentive constraint will be violated. Figure below shows the graph of revenue when α is 0.1.

If $\alpha > \frac{1}{4}$, we get that $\pi_1^* > \pi_2^*$. When $I \leq \pi_2^* = \frac{1}{4}\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$, both the feasibility and incentive constraints can be satisfied by π_1^* and π_2^*. Therefore, the revenue is $\pi_1^* + \pi_2^* - I = \frac{(2\alpha+1)^2}{4(4\alpha+1)} - I$, which is linear in I. When $\frac{1}{4}\frac{(2\alpha+1)^2}{(4\alpha+1)^2} < I \leq (\frac{2\alpha+1}{4\alpha+3})^2$, where $(\frac{2\alpha+1}{4\alpha+3})^2$ is the threshold value of I mentioned above, the entrepreneur should raise p_1 to

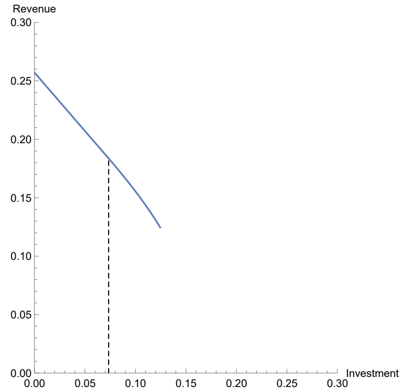


Figure 1.1: Entrepreneur's revenue when $\alpha = 0.1$

make the incentive constraint binding. Under such condition, $p_1 = (2\alpha + 1)\sqrt{I}$ and the revenue is equal to $p_1 - \frac{2}{2\alpha+1}p_1^2 = (2\alpha + 1)(\sqrt{I} - 2I)$, which is non-linear in I. When I exceeds $(\frac{2\alpha+1}{4\alpha+3})^2$, either feasibility constraint or incentive constraint will be violated. Here we show the graph of revenue when α is 0.8.

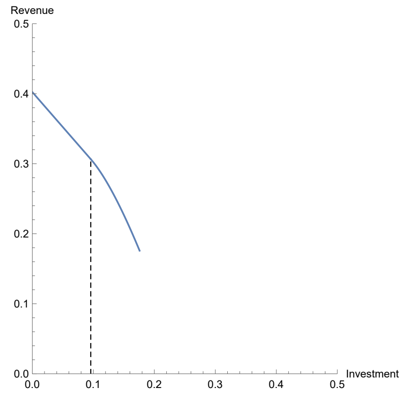


Figure 1.2: Entrepreneur's revenue when $\alpha = 0.8$

Omitted analysis on full information

In this section, we complete the analysis on the full information case with low-quality product. With the same method mentioned in Section 1.4.2.1, we analyze from $t = 2$. The entrepreneur has to maximize his profit at this stage, that is,

$$\max_{p_2^l} \; p_2^l(\bar{v}^l - \frac{p_2^l}{1-\Delta}) \tag{1.13}$$

where we can get $p_2^l = \frac{1-\Delta}{2}\bar{v^l}$. Then we come back to the first stage. Considering the consumer indifferent between pledging and waiting, he will get $(1-\Delta)(1+\alpha)\bar{v^l} - p_1^l$ if he decides to pledge. Otherwise, his payoff would be $(1-\Delta)\bar{v^l} - p_2^l$. By equalizing these two equations, that is,

$$(1-\Delta)(1+\alpha)\bar{v^l} - p_1^l = (1-\Delta)\bar{v^l} - p_2^l \tag{1.14}$$

we obtain that $\bar{v^l} = \frac{2p_1^l}{(1-\Delta)(2\alpha+1)}$ and $p_2^l = \frac{p_1^l}{2\alpha+1}$. Eventually, as the entrepreneur, he maximizes the total profits in two stages,

$$\max_{p_1^l} \; p_1^l(1-\bar{v^l}) + p_2^l(\bar{v^l} - \frac{p_2^l}{1-\Delta}). \tag{1.15}$$

Therefore we could easily calculate that $p_1^l = \frac{(1-\Delta)(2\alpha+1)^2}{8\alpha+2}$, $p_2^l = \frac{(1-\Delta)(2\alpha+1)}{8\alpha+2}$ and $\bar{v^l} = \frac{2\alpha+1}{4\alpha+1}$. Additionally, we can get $\pi_1^l = \alpha(1-\Delta)\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$, $\pi_2^l = \frac{1-\Delta}{4}\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$, and $\pi_l = \pi_1^l + \pi_2^l = \frac{1-\Delta}{4}\frac{(2\alpha+1)^2}{4\alpha+1}$.

Omitted analysis on symmetric uncertainty

In Section 1.4.2.2 we have derived that $p_1^s = p_1^* = \frac{(2\alpha+1)^2}{8\alpha+2}$ and $\bar{v}^s = \bar{v}^* = \frac{2\alpha+1}{4\alpha+1}$, so the profit in the first stage is $\alpha\frac{(2\alpha+1)^2}{(4\alpha+1)^2}$, which is the same as π_1^*. By plugging these results into p_2^h and p_2^l, we have that $p_2^h = \frac{1+\Delta}{2}\bar{v} = \frac{1+\Delta}{1+2\alpha}p_1^s$ and $p_2^l = \frac{1-\Delta}{2}\bar{v} = \frac{1-\Delta}{1+2\alpha}p_1^s$. When consumer's valuation is $(1+\Delta)v$, the profit in the second stage is $\frac{1+\Delta}{4}\bar{v}^2$. When consumer's valuation is $(1-\Delta)v$, the profit becomes $\frac{1-\Delta}{4}\bar{v}^2$. Therefore, the expected profit in the second stage $E(\pi_2^s) = \frac{\bar{v}^2}{4}$, which is identical to π_2^* in the benchmark case. Under such circumstance, the feasibility constraint means that $\pi_1^s \geq I$ and incentive constraint can be written as $E(\pi_2^s) \geq I$, both of which are identical to the benchmark case. Thus, the equilibrium outcome in the symmetric uncertainty case is the same as the benchmark model.

Next, we check the more general case in which there is probability β that the final product has high quality, and probability $1-\beta$ of low quality. With the backward induction, we start from the second stage and have that $p_2^h = \frac{1+\Delta}{2}\bar{v}$ and $p_2^l = \frac{1-\Delta}{2}\bar{v}$. Back to the first stage, considering the indifferent consumer $\bar{v}$, we have

$$\beta(1+\alpha)(1+\Delta)\bar{v} + (1-\beta)(1+\alpha)(1-\Delta)\bar{v} - p_1 = \beta[(1+\Delta)\bar{v} - p_2^h] + (1-\beta)[(1-\Delta)\bar{v} - p_2^l]$$

where we can get that $\bar{v} = \frac{2p_1}{(2\alpha+1)[1+(2\beta-1)\Delta]}$. In the beginning of the first stage, the entrepreneur chooses p_1 to maximize the expected profits

$$\max_{p_1}\ p_1(1-\bar{v}) + \beta p_2^h(\bar{v} - \frac{p_2^h}{1+\Delta}) + (1-\beta)p_2^l(\bar{v} - \frac{p_2^l}{1-\Delta}).$$

By inserting p_2^h and p_2^l we derive that $p_1^s = \frac{(2\alpha+1)^2}{8\alpha+2}[1+(2\beta-1)\Delta]$ and $\bar{v} = \frac{2\alpha+1}{4\alpha+1}$.

If $\beta < \frac{1}{2}$, $p_1^s < p_1^*$ and $\bar{v}$ doesn't change. Therefore, both π_1^s and $E(\pi_2^s)$ will decrease compared with π_1^* and π_2^*. To the contrary, if $\beta > \frac{1}{2}$, $p_1^s > p_1^*$ and both π_1^s and $E(\pi_2^s)$ will increase due to the less risk with the crowdfunding.

Proof of Corollary 1.1

This comes directly from the fact that $\bar{v^h} = \bar{v^l}$. As we have shown, if the entrepreneur lies, his expected profits will be $p_1^h(1-\bar{v^h}) + p_2^l(\bar{v^h} - \frac{p_2^l}{1-\Delta})$. To the contrary, if he reports truthfully, his expected profits would be $p_1^l(1-\bar{v^l}) + p_2^l(\bar{v^l} - \frac{p_2^l}{1-\Delta})$, where all elements are the same as the case if he deviates, except for p_1^l in the first term. Since p_1^l is smaller than p_1^h, the expected profits of truthfully reporting are lower.

Proof of Lemma 1.1

Let's first assume that low-quality entrepreneur deviates and mimics the behavior of high-quality entrepreneur, meaning that he sets p_{1s}^h in the first stage. With the same method as Equation (1.8) in the full information case, we can get $\bar{v}_s^h = \frac{2p_{1s}^h}{(1+\Delta)(2\alpha+1)}$. In the second stage, the quality reveals and he sets $p_{2s}^{l'}$ to maximize his profit among the residual demands, that is, $p_{2s}^{l'}(\bar{v}_s^h - \frac{p_{2s}^{l'}}{1-\Delta})$. By deriving the first order condition, we obtain that $p_{2s}^{l'} = \frac{1-\Delta}{2}\bar{v}_s^h = \frac{1-\Delta}{1+\Delta}\frac{p_{1s}^h}{2\alpha+1}$. Now, we can express the total profits of deviating $\pi_l^{'}$ as a function of p_{1s}^h,

$$\begin{aligned}
\pi_l^{'} &= p_{1s}^h(1-\bar{v}_s^h) + p_{2s}^{l'}(\bar{v}_s^h - \frac{p_{2s}^{l'}}{1-\Delta}) \\
&= p_{1s}^h - \frac{2}{(1+\Delta)(2\alpha+1)}{p_{1s}^h}^2 + \frac{1-\Delta}{(1+\Delta)^2(2\alpha+1)^2}{p_{1s}^h}^2 \\
&= -\frac{1+4\alpha+3\Delta+4\alpha\Delta}{(1+\Delta)^2(2\alpha+1)^2}[p_{1s}^h - \frac{(1+\Delta)^2(2\alpha+1)^2}{2(1+4\alpha+3\Delta+4\alpha\Delta)}]^2 + \frac{(1+\Delta)^2(2\alpha+1)^2}{4(1+4\alpha+3\Delta+4\alpha\Delta)}.
\end{aligned}$$

If, for the low-quality entrepreneur, $\pi_l^{'}$ is not larger than the full information profit,[19] he

[19] We have derived the profit under full information for the low-quality entrepreneur π_l in Equation (1.15), so it means that $\pi_l^{'} \le \pi_l$.

has no incentive to mimic the high-quality entrepreneur's strategy. By solving the inequality $\pi_l^{'} \leq \pi_l$, we have that when $p_{1s}^h \geq \frac{(1+\Delta)^2(2\alpha+1)^2}{2(1+4\alpha+3\Delta+4\alpha\Delta)} + \frac{(1+\Delta)(2\alpha+1)^2}{1+4\alpha+3\Delta+4\alpha\Delta}\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}$ or $p_{1s}^h \leq \frac{(1+\Delta)^2(2\alpha+1)^2}{2(1+4\alpha+3\Delta+4\alpha\Delta)} - \frac{(1+\Delta)(2\alpha+1)^2}{1+4\alpha+3\Delta+4\alpha\Delta}\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}$, separating equilibria exist. Moreover, the best separating equilibrium for the high-quality entrepreneur is to maximize his total profits. Thus, $p_{1s}^h = \frac{(1+\Delta)^2(2\alpha+1)^2}{2(1+4\alpha+3\Delta+4\alpha\Delta)} + \frac{(1+\Delta)(2\alpha+1)^2}{1+4\alpha+3\Delta+4\alpha\Delta}\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}$, and we can get $p_{2s}^h = \frac{(1+\Delta)(1-\Delta)(2\alpha+1)}{2(1+4\alpha+3\Delta+4\alpha\Delta)} + \frac{(1-\Delta)(2\alpha+1)}{1+4\alpha+3\Delta+4\alpha\Delta}\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}$ accordingly.

Proof of Lemma 1.2

We compare the profits in two stages separately. Denote π_{1s}^h, π_{2s}^h and π_s^h as the first-stage profit, second-stage profit and the total profits for the high-quality entrepreneur in the optimal separating equilibrium. π_1^h, π_2^h and π^h represent the respective profits under the full information. In the separating equilibrium we have that $\pi_{1s}^h = p_{1s}^h - \frac{2}{(1+\Delta)(2\alpha+1)}{p_{1s}^h}^2$, and under the full information $\pi_1^h = p_1^h - \frac{2}{(1+\Delta)(2\alpha+1)}{p_1^h}^2$, where p_{1s}^h and p_1^h stand for the first-stage prices in the optimal separating equilibrium and in the full information case. Thus, $\pi_{1s}^h - \pi_1^h = (p_{1s}^h - p_1^h)[1 - \frac{2}{(1+\Delta)(2\alpha+1)}(p_{1s}^h + p_1^h)]$. The first term $(p_{1s}^h - p_1^h)$ is positive since in the separating equilibrium the high-quality entrepreneur uses the higher first-stage price to signal the product quality. By inserting p_{1s}^h and p_1^h, the second term $1 - \frac{2}{(1+\Delta)(2\alpha+1)}(p_{1s}^h + p_1^h) = \frac{2\alpha}{4\alpha+1} - \frac{(1+\Delta)(2\alpha+1)}{1+4\alpha+3\Delta+4\alpha\Delta} - \frac{2(2\alpha+1)}{1+4\alpha+3\Delta+4\alpha\Delta}\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}} < 0$ because we can easily show that $\frac{2\alpha}{4\alpha+1} - \frac{(1+\Delta)(2\alpha+1)}{1+4\alpha+3\Delta+4\alpha\Delta} = -\frac{1+4\alpha+\Delta}{(4\alpha+1)(1+4\alpha+3\Delta+4\alpha\Delta)} < 0$. Therefore, $\pi_{1s}^h < \pi_1^h$ and the high-quality entrepreneur lowers his profit in the first stage to signal the product quality. On the other hand, $\pi_{2s}^h = \frac{{p_{1s}^h}^2}{(1+\Delta)(2\alpha+1)^2}$ and $\pi_2^h = \frac{{p_1^h}^2}{(1+\Delta)(2\alpha+1)^2}$. Thus, we have $\pi_{2s}^h > \pi_2^h$ since $p_{1s}^h > p_1^h$. Overall, in the optimal separating equilibrium, the total profits for the high-quality entrepreneur are reduced. To the contrary, the profit in the second stage increases, which lessens the risk of running away.

We can also derive the following result when making the further comparisons on the profits.

Corollary 1.2. *In the optimal separating equilibrium, the low-quality entrepreneur gets more profit in the first stage than the high-quality entrepreneur, but less profit in the second stage. Therefore, we have that* $\pi_{1s}^l > \pi_{1s}^h$ *and* $\pi_{2s}^l < \pi_{2s}^h$.

Proof. We denote π_s^h and π_s^l as the total profits for the high-quality entrepreneur and low-quality entrepreneur in the optimal separating equilibrium. Addtionally, we use $\pi_s^{l'}$ to represent the total profits when the low-quality entrepreneur deviates and pretends to be the high type in the first stage. In the optimal solution, we have that $\pi_s^l = \pi_s^{l'}$. We can write down these three profits in the following way:

$$\pi_s^l = \pi_{1s}^l + \pi_{2s}^l = p_1^l[1 - \frac{2p_1^l}{(1-\Delta)(2\alpha+1)}] + \frac{p_1^l}{2\alpha+1} \cdot \frac{p_1^l}{(1-\Delta)(2\alpha+1)}$$
$$\pi_s^{l'} = \pi_{1s}^{l'} + \pi_{2s}^{l'} = p_{1s}^h[1 - \frac{2p_{1s}^h}{(1+\Delta)(2\alpha+1)}] + \frac{1-\Delta}{1+\Delta} \cdot \frac{p_{1s}^h}{2\alpha+1} \cdot \frac{p_{1s}^h}{(1+\Delta)(2\alpha+1)}$$
$$\pi_s^h = \pi_{1s}^h + \pi_{2s}^h = p_{1s}^h[1 - \frac{2p_{1s}^h}{(1+\Delta)(2\alpha+1)}] + \frac{p_{1s}^h}{2\alpha+1} \cdot \frac{p_{1s}^h}{(1+\Delta)(2\alpha+1)}$$

Let $\frac{p_1^l}{2\alpha+1} \cdot \frac{p_1^l}{(1-\Delta)(2\alpha+1)} = A$ and $\frac{1-\Delta}{1+\Delta} \cdot \frac{p_{1s}^h}{2\alpha+1} \cdot \frac{p_{1s}^h}{(1+\Delta)(2\alpha+1)} = B$. Thus, $\frac{A}{B} = \frac{(1+\Delta)^2 {p_1^l}^2}{(1-\Delta)^2 {p_{1s}^h}^2}$. Under the full information we have shown that $\frac{{p_1^l}^2}{{p_1^h}^2} = \frac{(1-\Delta)^2}{(1+\Delta)^2}$. Additionally, in the separating equilibrium $p_{1s}^h > p_1^h$. Therefore, $\frac{{p_1^l}^2}{{p_{1s}^h}^2} < \frac{(1-\Delta)^2}{(1+\Delta)^2}$ and $\frac{A}{B} < 1$. So $A < B$ and $\pi_{2s}^l < \pi_{2s}^{l'} < \pi_{2s}^h$. Combining with the fact that $\pi_s^l = \pi_s^{l'}$, we get that $\pi_{1s}^l > \pi_{1s}^{l'} = \pi_{1s}^h$. This completes the proof that $\pi_{1s}^l > \pi_{1s}^h$ and $\pi_{2s}^l < \pi_{2s}^h$.

□

Omitted analysis on separating equilibrium

In this section, we compare the profits from two stages in the separating equilibrium. As we have shown, $\pi_{1s}^h = p_{1s}^h(1-\bar{v}_s^h) = p_{1s}^h - \frac{2}{(1+\Delta)(2\alpha+1)}{p_{1s}^h}^2$ and $\pi_{2s}^h = p_{2s}^h(\bar{v}_s^h - \frac{p_{2s}^h}{1+\Delta}) = \frac{1}{(1+\Delta)(2\alpha+1)^2}{p_{1s}^h}^2$. Thus, $\pi_{1s}^h - \pi_{2s}^h = p_{1s}^h[1 - \frac{2p_{1s}^h}{(1+\Delta)(2\alpha+1)} - \frac{p_{1s}^h}{(1+\Delta)(2\alpha+1)^2}]$. Plugging the optimal p_{1s}^h into it we can get that $\pi_{1s}^h - \pi_{2s}^h = \frac{p_{1s}^h}{2(1+4\alpha+3\Delta+4\alpha\Delta)}[4\alpha + 3\Delta + 4\alpha\Delta - 1 - (8\alpha+6)\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}]$.

When $\alpha \leq \frac{1}{4}$, $4\alpha - 1 \leq 0$ and therefore $4\alpha + 3\Delta + 4\alpha\Delta - 1 \leq 3\Delta + 4\alpha\Delta$. It is easy to verify that $3\Delta + 4\alpha\Delta < (8\alpha+6)\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}$. Thus, $\pi_{1s}^h - \pi_{2s}^h < 0$ under such circumstance, which matches the result in the main analysis. When $\alpha > \frac{1}{4}$, in order to determine the sign of $\pi_{1s}^h - \pi_{2s}^h$, we compare $4\alpha+3\Delta+4\alpha\Delta-1$ and $(8\alpha+6)\sqrt{\frac{2\alpha\Delta+2\alpha\Delta^2+\Delta^2}{4\alpha+1}}$, which is equivalent to check the values between $(4\alpha + 3\Delta + 4\alpha\Delta - 1)^2(4\alpha+1)$ and $(8\alpha+6)^2(2\alpha\Delta + 2\alpha\Delta^2 + \Delta^2)$. We can calculate that $(4\alpha + 3\Delta + 4\alpha\Delta - 1)^2(4\alpha+1) - (8\alpha+6)^2(2\alpha\Delta + 2\alpha\Delta^2 + \Delta^2) = (64\alpha^3 + 1) - (64\alpha^3\Delta^2 + 144\alpha^2\Delta^2 + 96\alpha^2\Delta + 108\alpha\Delta^2 + 80\alpha\Delta + 27\Delta^2 + 6\Delta + 4\alpha + 16\alpha^2)$. Based on the different values of α and Δ, π_{1s}^h can be either larger or smaller than π_{2s}^h. For instance, when Δ converges to 0, the equation above is equal to $64\alpha^3 + 1 - 16\alpha^2 - 4\alpha$. Then it's easy to show that $64\alpha^3 + 1 - 16\alpha^2 - 4\alpha$ is always positive when $\alpha > \frac{1}{4}$. So in this situation π_{1s}^h is larger than π_{2s}^h. However, if α converges to $\frac{1}{4}$, π_{1s}^h becomes smaller than π_{2s}^h. Overall, we can conclude that in the optimal separating equilibrium, if $\alpha \leq \frac{1}{4}$ we always have $\pi_{1s}^h < \pi_{2s}^h$. If $\alpha > \frac{1}{4}$, π_{1s}^h can be either larger or smaller than π_{2s}^h depending on the values of α and Δ.

Omitted analysis on pooling equilibrium

We have shown that when $p_1^h < p_{1s}^h$, there will be no separating equilibrium, since low-quality entrepreneur always has incentive to deviate to this strategy. In this part we check the pooling equilibrium under such circumstance.

In the first stage consumers can't tell the quality from the pre-ordering price, therefore they have a modified belief that there is a probability of $\frac{1}{2}$ with the high-quality product and another probability of $\frac{1}{2}$ with low-quality product. For consumers, this situation is the same as the case with symmetric uncertainty. By pledging in the first stage, the expected payoff will be

$$\frac{1}{2}(1+\alpha)(1+\Delta)\bar{v} + \frac{1}{2}(1+\alpha)(1-\Delta)\bar{v} - p_1.$$

If this consumer decides to wait and buy in the second stage, his expected payoff will be

$$\frac{1}{2}[(1+\Delta)\bar{v} - p_2^h] + \frac{1}{2}[(1-\Delta)\bar{v} - p_2^l]$$

where p_2^h and p_2^l are derived from the second stage when the quality has revealed, by maximizing $p_2^h(\bar{v} - \frac{p_2^h}{1+\Delta})$ and $p_2^l(\bar{v} - \frac{p_2^l}{1-\Delta})$ respectively. Inserting $p_2^h = \frac{1+\Delta}{2}\bar{v}$ and $p_2^l = \frac{1-\Delta}{2}\bar{v}$ into the two expected payoffs above and equalizing them, we can get that $\bar{v} = \frac{2p_1}{2\alpha+1}$. In a pooling equilibrium, it is the high-quality entrepreneur who decides the price to maximize his expected profits, that is,

$$\max_{p_{1p}} \; p_{1p}(1-\bar{v}) + p_2^h(\bar{v} - \frac{p_2^h}{1+\Delta}). \tag{1.16}$$

Finally we can get that $p_{1p} = \frac{(2\alpha+1)^2}{8\alpha+2-2\Delta}$ and $\bar{v} = \frac{2\alpha+1}{4\alpha+1-\Delta}$ as the optimal pooling equilibrium.[20]

The feasibility of crowdfunding will be affected in pooling equilibrium, since consumers cannot identify the product quality. Concerning the high-quality entrepreneur, this can be verified by equalizing the first- and second-stage profit in Equation 1.16, that is $p_{1p}(1-\bar{v}) = p_2^h(\bar{v} - \frac{p_2^h}{1+\Delta})$. By inserting $\bar{v}$ and p_2^h derived above, we can get that $p_{1p} = \frac{(2\alpha+1)^2}{4\alpha+\Delta+3}$. Thus, the largest possible investment requirement in the pooling case is

[20] Please note that $1-\bar{v}$ should be positive, in order to ensure that the profit in the first stage is positive and thereby the crowdfunding is feasible. Thus, we should make a further assumption that $\Delta < 2\alpha$ considering this scenario.

$\frac{(1+\Delta)(2\alpha+1)^2}{(4\alpha+\Delta+3)^2}$. Following the same method in the benchmark case, we can calculate that the largest possible investment requirement under the full information is $\frac{(1+\Delta)(2\alpha+1)^2}{(4\alpha+3)^2}$ for the high-quality entrepreneur.

Overall, it shows that for the high-quality entrepreneur, less projects will be feasible in the pooling equilibrium compared with the full information case. Similarly, we can demonstrate that the condition for the low-quality entrepreneur becomes better off in the pooling case.

Omitted analysis on commitment

If the entrepreneur could make the commitment in the benchmark model, he would choose (p_1, p_2) simultaneously to maximize $p_1(1-\frac{p_1-p_2}{\alpha})+p_2(\frac{p_1-p_2}{\alpha}-p_2)-I$, where $\frac{p_1-p_2}{\alpha}$ represents the indifferent consumer between pledging in the first stage and waiting until the retail market. It results in $p_1=\frac{1}{2}+\frac{\alpha}{2}$ and $p_2=\frac{1}{2}$, and the indifferent consumer $\bar{v}=\frac{p_1-p_2}{\alpha}$ is located at $\frac{1}{2}$. This means that no one buys from the retail market, which matches with the existing theory about intertemporal price discrimination on durable goods. If we compare this result with the benchmark model, it is easy to see that the total profits increase at the cost of the potential moral hazard problem. Under such circumstance, the entrepreneur could raise the most money from crowdfunding. On the other hand, this is the worst scenario for consumers, due to the less surplus and the most severe moral hazard problem.

We can calculate consumer surplus in the non-commitment and commitment cases,

$$CS^{nc}=(1+\alpha)\frac{1+\bar{v}^*}{2}(1-\bar{v}^*)+\frac{\bar{v}^*+p_2^*}{2}(\bar{v}^*-p_2^*)-p_1^*(1-\bar{v}^*)-p_2^*(\bar{v}^*-p_2^*)=\frac{16\alpha^3+36\alpha^2+12\alpha+1}{8(4\alpha+1)^2}$$

$$CS^c=(1+\alpha)\frac{1+\bar{v}}{2}(1-\bar{v})-p_1(1-\bar{v})=\frac{\alpha+1}{8}$$

where we have $CS^{nc}>CS^c$. Producer surplus can be computed as

$$PS^{nc}=p_1^*(1-\bar{v}^*)+p_2^*(\bar{v}^*-p_2^*)-I=\frac{(2\alpha+1)^2}{4(4\alpha+1)}-I$$

$$PS^c=p_1(1-\bar{v})-I=\frac{\alpha+1}{4}-I$$

where we get $PS^{nc}<PS^c$. Thus, we have the following social surplus,

$$W^{nc}=CS^{nc}+PS^{nc}=\frac{48\alpha^3+76\alpha^2+28\alpha+3}{8(4\alpha+1)^2}-I$$

$$W^c = CS^c + PS^c = \frac{48\alpha^3 + 72\alpha^2 + 27\alpha + 3}{8(4\alpha+1)^2} - I.$$

It shows $W^{nc} > W^c$, therefore the social planner has the same incentive with consumers that prefers the non-commitment case.

Now we prove that when taking the feasibility and incentive constraints into account, crowdfunding with commitment is better than that without commitment. First, it is straightforward that in the crowdfunding with commitment, the entrepreneur can replicate the optimal strategies that we mentioned above to make himself at least indifferent to the case where commitment is not allowed. Next, we need the following lemma to help us figure out the strategy, which makes the entrepreneur strictly better off.

Lemma 1.3. *When (p_1, p_2) satisfies $p_1 = (2\alpha+1)p_2$ as the optimal strategy in the benchmark model, the entrepreneur can increase the total profits by keeping p_1 the same as the optimal result while raising p_2 by a small amount.*

Proof. In the benchmark case, $\pi_1 = p_1(1 - \frac{p_1-p_2}{\alpha})$ and $\pi_2 = p_2(\frac{p_1-p_2}{\alpha} - p_2)$. When we increase p_2 to p_2' by the amount of γ, the new profits are $\pi_1' = p_1(1 - \frac{p_1-p_2'}{\alpha})$ and $\pi_2' = p_2'(\frac{p_1-p_2'}{\alpha} - p_2')$ where $p_2' = p_2 + \gamma$. We can calculate that the change of the profit in the first stage is $\Delta\pi_1 = \frac{p_1}{\alpha}\gamma$, the change of the profit in the second stage is $\Delta\pi_2 = -\gamma[\frac{p_2}{\alpha} + \frac{\alpha+1}{\alpha}\gamma]$, and the change of the total profits is $\Delta\pi = \gamma[\frac{2\alpha+1}{4\alpha+1} - \frac{\alpha+1}{\alpha}\gamma]$. When γ is positive and converging to zero, we have that $\Delta\pi_1 > 0$, $\Delta\pi_2 < 0$ and $\Delta\pi > 0$. □

Lemma 1.3 shows a profitable deviation for the entrepreneur. Now we need to confirm that this deviation also satisfies the feasibility and incentive constraints. Based on the result in the benchmark model, we discuss two cases separately. Let us first consider the case with $\alpha < \frac{1}{4}$. Recall that when $I \leq \pi_1^*$, (p_1^*, p_2^*) constitutes the optimal strategy. The entrepreneur could commit to $(p_1^*, p_2^* + \gamma)$ where $\gamma > 0$, in order to get the higher profits and make both constraints still satisfied. When $I > \pi_1^*$, as analyzed in the benchmark case, the entrepreneur has to lower p_1 to make the feasibility constraint binding, that is $\pi_1 = I$. Now with the proposed deviation, π_1 is increased compared with the benchmark model and therefore the crowdfunding is also feasibile. With a similar method we can verify the profitable deviation in the case with $\alpha > \frac{1}{4}$. Overall we demonstrate that the ability to commit makes the entrepreneur strictly better off.

Chapter 2

Behavior-based Price Discrimination under Endogenous Privacy

Based on Heiny, Li and Tolksdorf (2020).

2.1 Introduction

With an increased capability to process big data and the passing of EU's General Data Protection Regulation (GDPR), behavior-based price discrimination (BBPD) and consumer privacy have become a hot topic. There exists evidence that retailers use consumers' data to price discriminate between them. Mikians et al. (2012, 2013) conduct online field experiments and find evidence for price and search discrimination in e-commerce based on geographical location and consumer's budget. Consumers' data are also used for personalizing advertisements. Tucker (2012) summarizes a wide literature on targeted advertising which is feasible due to a certain degree of privacy intrusion. The first major web experiment of behavior-based price discrimination was conducted by Amazon as early as 2000 (Streitfeld, 2000). The company discriminated between consumers based on the number of previous purchases at Amazon. Since then using consumers' private data for behavior-based pricing has become a common practice in online retailing.

Since Amazon's experiment a lot has changed in the field of data protection and privacy. Particularly, the passing of the GDPR in May 2018 was a major breakthrough for privacy protection. In accordance with the regulation, consumers can now decide whether to allow websites to access their personal information contained in cookies (Parliament and Council of the European Union, 2016). Cookies are placed by websites to track and record information about previous visits and online activities.[1] The collected data are used by online retailers to make personalized offers in line with behavior-based pricing. The GDPR gives consumers control over their personal information by having a choice to opt-out. By refusing cookies, consumers stay anonymous and cannot be identified (as previous customers). Conversely, when consumers allow a firm to access their cookies, they can be identified and targeted with customized prices. In economic terms, this choice to opt-out means consumers can act strategically. The GDPR is designed to make consumers' data exclusively accessible to one retailer at a time. If that retailer plans on distributing consumers' data to third parties, consumers have to explicitly consent.[2]

In this chapter, we develop a model of behavior-based pricing à la Fudenberg and Tirole (2000) with an endogenous privacy decision of consumers. We compare consumers' and firms' behavior in two distinct data environments. First, based on the design of the GDPR, we look at an exclusive data environment, i.e., only the seller a consumer bought from receives consumers' information. Second, we propose an open data environment, i.e., data sharing among firms is mandated. Even though there is no mandated data sharing policy among businesses in place currently, the European Commission explores the idea of business-to-business data sharing in a recent strategy proposal (European Commission, 2020). For the public sector, there is already an 'Open Data Directive' in place to facilitate data sharing among institutions. We address the following questions: how do consumers react to behavior-based pricing when their privacy choice is endogenous? How do sellers change their pricing strategy? And how does consumers' and firms' behavior differ in the two data environments?

In the first part of this chapter, we solve our theoretical model for pure-strategy equilibria that determine consumers' strategy concerning their privacy and firms' price setting. In the second part, we implement the model as a market game in a laboratory experiment with human subjects taking the role of consumers and sellers. Considering this is a real-world problem with practical implications for policy, we want to explore whether subjects follow our predicted strategies and more specifically whether consumers act rationally in their privacy decision. The experimental literature shows that this is not necessarily the case when privacy is concerned (Acquisti et al., 2016, Schudy and Utikal, 2017), which we can confirm with our experimental data as well.

[1] Throughout this chapter we use the term "cookie" to refer to information about past purchases.

[2] It is common among retailers to give consumers' information to a data processor for an analysis of the information.

The GDPR provides an easy way for consumers to protect their privacy. We help understand how the privacy policy of the EU altered consumers' and consequently firms' behavior. We also contribute to the debate about a policy that mandates data sharing among firms and extends the 'Open Data Directive' to the private sector.

Following Fudenberg and Tirole (2000), we build on the Hotelling (1929) linear city model with two competing firms and a continuum of consumers. We consider a two-period game, where a consumer buys one unit of a non-durable product in each period from one of the firms. In the first period, firms set identical prices for all consumers with no information about consumers' preferences. Consumers then decide from which firm to buy and whether to accept the use of cookies. In the second period, based on consumers' strategy, firms can set different prices and consumers then, again, decide where to buy. In an open data environment, firms share the obtained information with each other, for example via cookie matching (Ghosh et al., 2015). In the exclusive data environment, given consumers accept cookies, only a consumer's supplier can access the private information.

We conduct a laboratory experiment, because this gives us full control over the data environment and consumers' privacy choices are observable. The stylized market we design closely resembles our theoretical model. One distinguishing feature is that subjects in the experiment play the market game for 20 consecutive rounds. We focus subjects attention on the strategic interaction, while also controlling for their stance on privacy issues and their cognitive capability for strategic interactions.

In the theoretical analysis, we find that the optimal pricing strategy is a mixture of uniform pricing and standard behavior-based pricing depending on the equilibrium strategies of consumers. In the open data environment, when consumers' data are available to both competitors, consumers in equilibrium choose to reveal their data, in order to increase competition between sellers.[3] Firms use the data to price discriminate between loyal consumers and consumers, who previously purchased from the competitor. We can confirm this result with experimental evidence from the open data treatment. Consumers predominantly allow tracking of their past purchase, which gives sellers the chance to use behavior-based pricing. We observe that sellers use poaching prices as reward for accepting cookies that are lower than loyalty or anonymous prices.

In the exclusive data environment, when consumers' data are only available to the respective firm they bought from, consumers are individually best off by maintaining their privacy. To our knowledge, this is a novel finding in competitive settings with

[3] Ali et al. (2019) and Casadesus-Masanell and Hervas-Drane (2015) also support this result in their theoretical models.

privacy decisions. Clearly, firms can only use uniform prices in such a scenario. Our experimental results diverge from this theoretical result. The evidence exhibits a high rate of accepting cookies even in the exclusive treatment, however, there is a downwards trend in the cookies sharing rate over time. Even though a lot of consumers share their data, we only observe poaching efforts in later rounds of the experiment and to a much smaller extent compared to the open data environment.

From a theoretical point of view, social welfare is maximal in the exclusive data environment, even though it hurts consumers. In equilibrium all consumers choose to be anonymous and, hence, firms set uniform prices. In this equilibrium consumers do not switch between firms. On the other hand, in equilibrium consumer welfare is larger in the open data environment, because consumers benefit from poaching offers. The experimental results exhibit no significant difference in social welfare between treatments. Subjects are not fully rational in their privacy choice such that inefficient switching occurs in both treatments. Hence, our experimental analysis reveals that an open data environment can be an option to enhance competition without incurring a loss in total welfare. The welfare analyses demonstrate that there needs to be a discussion about whether mandated data sharing among firms should be implemented as a policy. Our theoretical and experimental results show that mandated data sharing among firms leads consumers to share more data. A policy maker, therefore, faces a trade-off between privacy protection and efficiency when establishing the relevant policy.

2.2 Related Literature

This chapter is related to a set of articles in the theoretical and the experimental literature. In the theoretical literature there are a few papers which also deal with endogeneous privacy and price discrimination (Acquisti and Varian, 2005, Conitzer et al., 2012, Belleflamme and Vergote, 2016, Ali et al., 2019).

The two papers closest to our research are Conitzer et al. (2012) and Ali et al. (2019). Conitzer et al. (2012) study a monopoly with an outside option and consumers can choose to let the monopolist track their purchases. They find that under free anonymization all consumers choose to do so, which gives the monopolist the highest payoff. Importantly, the introduction of competition raises issues in the handling of the information structure, which leads to our separation of the open and exclusive data environments. As in this chapter, consumers have an endogenous privacy choice. However, Conitzer et al. (2012) do not study a competitive situation of behavior-based pricing, where the strategic action of consumers has different implications for pricing. Our focus is on consumers' privacy choice for different data environments, in which we diverge from the theoretical analysis of Conitzer et al. (2012). Ali et al. (2019) study how complete consumer control affects personalized pricing in a monopoly as well as under competition. They focus on comparing disclosure channels and analyze consumers sharing rich and simple evidence about their types. In our model we only look at a dichotomous disclosure technique, tracking versus no tracking, but support the same result, that voluntary disclosure amplifies competition.

Colombo (2016) considers a setup of incomplete information sharing in a duopoly case similar to our exclusive data environment (in Section 2.4.3). Colombo uses a fixed parameter as share of anonymous consumers and does not consider consumers' endogenous privacy choice. The main point of our study, however, is to analyze the strategic decision of consumers. Other papers that are also concerned with price discrimination and exogenous privacy are Liu and Serfes (2004) and Esteves (2014).

Taylor (2004) and Montes et al. (2018) extend the idea of price discrimination and privacy to include a data broker. Montes et al. (2018) consider a duopoly with a costly privacy choice for consumers. They focus on a data broker who sells consumers' data to competing firms. One of their main results is that information is usually only sold to one of the firms. This result is what we capture in our exclusive data environment where we observe a higher producer surplus than in the open data envrionment.

Other papers expand the literature by looking at platform markets and consumer data (Casadesus-Masanell and Hervas-Drane, 2015, Braghieri, 2019). In Casadesus-Masanell and Hervas-Drane (2015) sellers derive profit from buyers' purchases as well as from a secondary market where consumers' information can be disclosed. The decision of consumers to reveal information plays a twofold role here. The authors find that the profit maximizing strategy of firms is to focus on the consumer market. Furthermore, they observe that under competition there is a lower level of disclosure in the market.

Extensive reviews of the literature on behavior-based price discrimination in general and the economics of privacy can be found in Armstrong (2006*b*), Fudenberg and Villas-Boas (2006), Esteves et al. (2009), Acquisti et al. (2016).

The analysis of behavior-based pricing under endogenous privacy in the experiment relates to two branches of experimental literature. Firstly, the basic structure and procedure are related to spatial competition experiments. We extend the existing literature on BBPD and spatial competition with location choice experiments. BBPD experiments have been conducted by Brokesova et al. (2014) and Mahmood (2014). Brokesova et al. (2014) computerize the buyer's side, which we do not. Mahmood (2014) only considers two fixed locations for buyers, whereby the experimental market rather resembles a Bertrand market with differentiated products than a spatial competition.

We employ a BBPD experiment similar to those two but introduce features from spatial competition with location choice experiments by Camacho-Cuena et al. (2005) and Barreda-Tarrazona et al. (2011), which is how we transform the theoretical setup into an experimental setup with treatments corresponding to the two data environments.

Secondly, we introduce privacy and data sharing elements. Similar issues have been studied before, but to our knowledge not in the context of an explicit market experiment. Acquisti et al. (2013) identify a considerable gap between willingness to accept disclosure of private information and willingness to pay for the protection of private information. To alleviate this issue we renounce enforcing a default option on privacy, assuming disclosure and protection are both costless. Beresford et al. (2012) and Preibusch et al. (2013) find that subjects have a remarkably low willingness-to-accept for giving up their privacy and are not acting on their stated privacy decisions when protection of privacy is costless. This finding contrasts Tsai et al. (2011) who find that subjects act on websites' certified privacy protection qualities when shopping online. They suggest that subjects might in fact be willing to pay premiums for privacy protection. Despite an overall high rate of information disclosure, we find that privacy considerate subjects[4] share less information under the exclusive data treatment, while there is no such effect in the open data treatment.

Schudy and Utikal (2017) find that subjects' willingness to share personal information decreases when the number of recipients of said information increases. Between our open and exclusive data treatments the number of recipients varies. In support of their findings, we observe a higher willingness to share information in the early rounds of the exclusive data treatment. Feri et al. (2016) explore in a lab setting how privacy disclosure is affected by risks of privacy shock in the form of data breaches. They find that only those consumers, who regard their information as sensitive, demonstrate an effect on information disclosure under different likelihoods of data breaches.

We contribute to the literature by focusing on consumers' endogenous privacy decisions in competitive markets under a set of different information schemes. Combining a theoretical model with an experiment is a novel approach to answer our research questions.

2.3 Model

We consider a setup following Hotelling (1929), where a line segment of length $\bar{\theta}$ spans a product characteristic space. Along the line, consumers are uniformly distributed with a density of $\bar{\theta}^{-1}$, i.e., we assume a consumer mass of 1. The location of a consumer is

[4]We depict those subjects who give more mixed responses in our privacy survey as *considerate*.

private knowledge to them and denoted by $\theta \in [0, \bar{\theta}]$, such that θ serves as a consumer's preferred variety of a good.

There are two firms each producing a variant of the same good at constant marginal costs normalized to zero; fixed costs are neglected. Firm A is located at the left end of the line segment, while firm B is placed at the right end. The firms compete for two periods, $t = 1, 2$.

In each period, consumers buy one unit of the good either from firm A or B, i.e., we assume that the product's valuation is large enough to make sure each consumer buys one unit in each period. No outside option is available. Considering a consumer located at $\hat{\theta}$, their utility is given by $U_A = v - p_a - \hat{\theta}$ or $U_B = v - p_b - (\bar{\theta} - \hat{\theta})$, depending on their purchasing decision. We assume consumers' unit transportation cost to be normalized to one. Consumers' valuations, v, are the same over time for all consumers. Their rationale is to maximize their utility. We do not take discounting into account.

On top of the purchasing decisions, consumers also decide whether to accept the use of cookies, $q \in \{0, 1\}$, in the first period. We use cookies as proxy for a consumer's purchasing history, which is revealed to a company if $q = 1$. In that case, a firm is able to identify a buyer from period $t = 1$ and can thus set a different price in the upcoming period. Consumers have the option to act strategically with regards to revealing information. In the literature, it is often assumed that generating privacy involves some costs (Conitzer et al., 2012, Montes et al., 2018). However, Loertscher and Marx (2020) show that the act of providing data can even be costly for consumers. We refrain from such an assumption to keep the theoretical predictions clean from any cost effects and not impose an implicit privacy concern on subjects in the experiment.[5]

In the first period, competing firms set prices $\mathbf{p_1} = (p_1^A, p_1^B)$. In the second period, pricing is more involved. Depending on the preceding cookie choice of consumers, there is a share λ of anonymous customers who forbade the use of their cookies and a share $1 - \lambda$ of identifiable customers. These shares are derived from the aggregation of consumers' choices regarding their cookies.[6]

Given the cookie choice of a consumer, we differentiate between two data environments. The data environments differ with regards to the number of recipients of consumers' data. Data contained in cookies are either shared between firms, which we call *open data environment* or firms keep them privately-held, which we call *exclusive data environment*.

In the open data environment, accepting the use of cookies means that both firms

[5] GDPR actually allows the easier way to opt in or opt out, which is in accordance with our assumption.

[6] An additional modeling assumption is that we use a common λ for consumers of both firms, i.e., identified and unidentified consumers of each firm are not countable. For example, due to market research firms could have a general understanding of how many consumers anonymize.

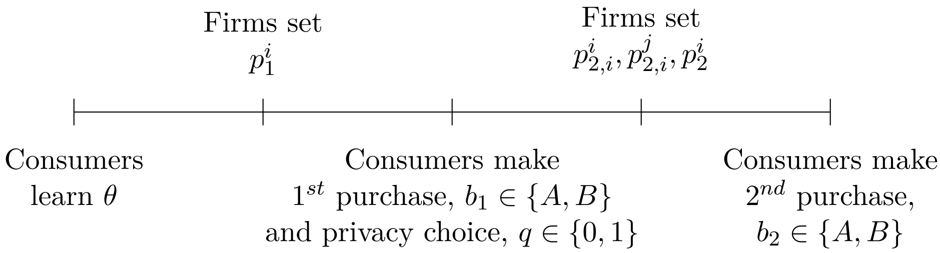


Figure 2.1: Timeline

can access the information about a consumer's past purchase (contained in the cookie). In the exclusive data environment, accepting the use of cookies means that only the firm a consumer has bought from can access information about a consumer's past purchase (contained in the cookie).

In the open data environment where both firms can target the competitors' consumers, each distinguishes three prices in the second period: $p_{2,i}^i$, is a loyalty price for consumers who bought from firm i in the first period and decide to buy from the same firm in the second period; $p_{2,i}^j$, is a poaching price for identifiable consumers who bought from i in the first period and j in the second period; and p_2^i, is an anonymous price for consumers who belong to the share λ, where $i, j = A, B$ and $i \neq j$. The idea of poaching consumers was first explored in Fudenberg and Tirole (2000).

In Section 2.4.3, we diverge from the open data environment and assume that only firms that consumers have bought from in the first period can learn about the purchasing history. This alters the pricing strategy considering that firms can no longer set a poaching price $p_{2,i}^j$, since the information needed is not available to them.

Figure 2.1 depicts the timing of the game. At the beginning, each consumer privately learns their type θ_i. Then in the first period, firms set prices $\mathbf{p_1}$. Afterward, consumers simultaneously make their purchasing decision, $b_1 \in \{A, B\}$ and their cookie choice, $q \in \{0, 1\}$. In the second period, firms set prices $\mathbf{p_2} = (p_{2,i}^i, p_2^i, p_{2,i}^j)$. At the end of the second period consumers again choose to buy from A or B. Finally, consumers receive their utilities and firms earn profits.

We solve for perfect Bayesian Nash equilibrium (PBE). In this context a PBE comprises firm's and consumers' strategy. Firms set first- and second-period prices given the sets for each period.[7] They also form beliefs about consumers' locations given their cookie choices, dependent on their type. Consumers make choices concerning first-period

[7]The strategy should also contain second-period prices if firms had set different prices in the first period. This is omitted here for simplicity.

purchase and cookies as functions of the location and p_1^i, as well as, second-period purchase as function of location and second-period prices.

2.4 Equilibrium Analysis

In this section, we determine the theoretical equilibrium of the two-period game. First, we discuss a benchmark model, where privacy is exogenous. We check two scenarios in which full privacy or no privacy is given, in order to make a comparison with the cases we analyze later where privacy is endogenous. Next, we consider the two aforementioned distinct information settings: firstly, we assume that information about previous purchases is given to both firms, irrespective of the actual buying decision of consumers, i.e., the *open data environment*. Secondly, we analyze the *exclusive data environment*, where only the firm that a consumer has bought from in $t = 1$ can access consumers' cookies.

Due to intractability of a more general specification, we focus on two types of beliefs concerning consumers' cookie setting behavior. The first specification resembles *pooling* in that we consider a vertical segmentation of consumers. That is, independent of their type consumers all make the same cookie choice. Firms observe the share of anonymous consumers and presume this to be identical to the probability to hide information for each consumer, independent of their type. The corresponding beliefs are $\hat{q}(\theta) = \lambda$. Under these specifications, we find that consumers fully disclose their data in the open data environment and fully hide in the exclusive data environment. By means of the intuitive criterion, we proceed to show that these *pooling* beliefs can be sustained off-path in support of our analysis. Though we cannot prove uniqueness amongst all potential candidates, we exclude another prominent class, which we refer to as *separating* cases. The *separating* cases correspond to an extreme segmentation of consumers, where we allow for any number of pure-strategy segments of arbitrary length. Within each segment all consumers either fully disclose or fully hide their information.

2.4.1 Benchmark

As a benchmark, we solve our dynamic model for an exogenously given privacy choice such that there are two cases: full privacy and no privacy. The results demonstrated here are drawn from the literature. In our analysis we show that even under endogenous privacy we can duplicate these results.

In case of full privacy, the two periods are completely independent. Firms do not obtain any information about consumers and cannot price discriminate. Thus, each firm sets a unique price in each period and tries to maximize the profit from that period. As a best response, each firm chooses the same prices $\bar{\theta}$ in both periods which corresponds to the results of Hotelling (1929).

In the case with no privacy, firms automatically access cookies of their own consumers and thereby recognize old customers.[8] The model becomes a standard behavior-based price discrimination introduced by (Fudenberg and Tirole, 2000) and Fudenberg and Villas-Boas (2006). In $t = 1$, competing firms set prices $\mathbf{p_1} = (p_1^A, p_1^B)$ and consumers decide where to buy. In $t = 2$, firms can identify all who have bought from them in the first period and set loyalty prices $p_{2,i}^i$ to these identified consumers. Meanwhile, for those who have bought from the competing firm in the first period, they face the anonymous prices p_2^i. From the conclusion in Fudenberg and Tirole (2000), we get that in the first period both firms will set the prices at $\frac{4}{3}\bar{\theta}$ and in the second period $p_{2,i}^i = \frac{2}{3}\bar{\theta}$ for the identified customers and $p_2^i = \frac{1}{3}\bar{\theta}$ for the unidentified customers.

2.4.2 Open Data Environment

In the open data environment, both firms receive information about a consumer's previous purchase given the consumer decides to grant access to their cookie. We begin our analysis under the *pooling* assumption. Firms hold the belief that $\hat{q}(\theta) = \lambda,\ \forall\theta$. The observed share of anonymous consumers λ corresponds to the probability that each consumer hides their information. Analogously, the counter-probability, $1 - \lambda$, corresponds to the probability that each consumer shares their information.

Consumers who did not let firms access their cookies in the first period are anonymous to both firms and are treated as new customers. Therefore, they face prices p_2^A (p_2^B) from firm A (firm B) in the second period. Consumers who revealed information about the purchase in the first period can be recognized by firms and thus are offered different prices in the second period. $p_{2,A}^A$ ($p_{2,A}^B$) are the prices offered by firm A (firm B) for those who bought from firm A in the first period, and $p_{2,B}^B$ ($p_{2,B}^A$) are provided by firm B (firm A) for those who bought from firm B in the first period.

The firms' beliefs about consumers' locations correspond to a segmentation of the Hotelling line. Starting from the second period, we divide consumers by their privacy choice into identifiable and anonymous consumers and can therefore consider two separate Hotelling lines. On the anonymous consumers' line, there is a mass of λ uniformly distributed consumers. Given the anonymous prices p_2^A and p_2^B, there exists a marginal consumer θ_2 who is indifferent between buying from firm A and firm B. θ_2 is determined

[8]With respect to the informational disclosure, this corresponds to our exclusive data environment.

by $v - p_2^A - \theta_2 = v - p_2^B - (\bar{\theta} - \theta_2)$ as

$$\theta_2 = \frac{\bar{\theta}}{2} + \frac{p_2^B - p_2^A}{2}.$$

Therefore, consumers with location $\theta \in [0, \theta_2)$ buy from firm A in the second period given they anonymized. Similarly, consumers with $\theta \in (\theta_2, \bar{\theta}]$ buy from firm B.

The other line has a mass of $1-\lambda$ uniformly distributed consumers who revealed their data in the first period. They are confronted with behavior-based price discrimination. Among the mass of $1-\lambda$ consumers, those who bought from firm A in the first period are given two prices in the second period: $p_{2,A}^A$ as loyalty price set by firm A and $p_{2,A}^B$ as a poaching price from firm B. Similarly, consumers who bought from firm B in the first period also face two prices now, $p_{2,B}^B$ as loyalty price from firm B, and $p_{2,B}^A$ as a poaching price from firm A.

On A's turf there is a marginal consumer indifferent between buying from firm A at $p_{2,A}^A$ and buying from firm B at $p_{2,A}^B$. They are characterized by

$$\theta_2^A = \frac{\bar{\theta}}{2} + \frac{p_{2,A}^B - p_{2,A}^A}{2}.$$

Accordingly, on B's turf the marginal customer θ_2^B is determined by

$$\theta_2^B = \frac{\bar{\theta}}{2} + \frac{p_{2,B}^B - p_{2,B}^A}{2}.$$

Identified consumers with $\theta \in [0, \theta_2^A)$ and $\theta \in (\theta_2^B, \bar{\theta}]$ are loyal to their first-period sellers. Whereas, consumers located at $\theta \in (\theta_2^A, \theta_1)$ and $\theta \in (\theta_1, \theta_2^B)$ are poached by the competitor firm, where θ_1 denotes the first-period marginal consumer who is indifferent between buying from A and B. Figure 2.2 depicts the consumer shares and respective pricing by spanning a rectangle over both lines connected vertically through the share λ under the belief $\hat{q}(\theta) = \lambda$.

The maximization problem of the firms concerning the anonymous consumers is given by:

$$\max_{p_2^A} \quad \lambda p_2^A \left[\frac{\bar{\theta}}{2} + \frac{p_2^B - p_2^A}{2}\right]$$
$$\max_{p_2^B} \quad \lambda p_2^B \left[\bar{\theta} - (\frac{\bar{\theta}}{2} + \frac{p_2^B - p_2^A}{2})\right].$$

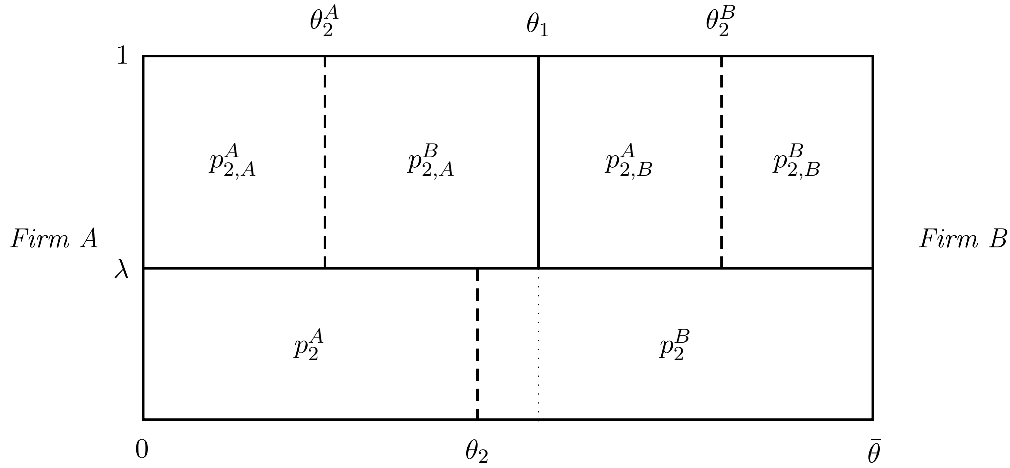


Figure 2.2: Customer segments under open data in $t = 2$

We can derive the first-order conditions as

$$\frac{\bar{\theta}}{2} + \frac{p_2^B}{2} - p_2^A = 0 \qquad \frac{\bar{\theta}}{2} - p_2^B + \frac{p_2^A}{2} = 0$$

and obtain the anonymous prices $p_2^A = p_2^B = \bar{\theta}$ and the marginal consumer $\theta_2 = \frac{\bar{\theta}}{2}$.

Among the identified consumers with mass $1 - \lambda$, we have the following maximization problems:

$$\max_{p_{2,A}^A, p_{2,B}^A} \quad (1-\lambda)\left[p_{2,A}^A \theta_2^A + p_{2,B}^A(\theta_2^B - \theta_1)\right]$$
$$\max_{p_{2,B}^B, p_{2,A}^B} \quad (1-\lambda)\left[p_{2,B}^B(\bar{\theta} - \theta_2^B) + p_{2,A}^B(\theta_1 - \theta_2^A)\right].$$

By plugging θ_2^A and θ_2^B into these two equations, we can solve for the prices.

Lemma 2.1. *The set of prices in the second period depend on $\theta_1(\mathbf{p_1})$ and the parameter $\bar{\theta}$.*

If $\frac{1}{4}\bar{\theta} \leq \theta_1 \leq \frac{3}{4}\bar{\theta}$, the equilibrium prices are:

$$p_2^A = \bar{\theta} \qquad p_{2,A}^A = \tfrac{1}{3}(2\theta_1 + \bar{\theta}) \qquad p_{2,B}^A = \tfrac{1}{3}(3\bar{\theta} - 4\theta_1)$$
$$p_2^B = \bar{\theta} \qquad p_{2,B}^B = \tfrac{1}{3}(3\bar{\theta} - 2\theta_1) \qquad p_{2,A}^B = \tfrac{1}{3}(4\theta_1 - \bar{\theta})$$

If $\theta_1 < \frac{1}{4}\bar{\theta}$, the equilibrium prices are:[9]

[9]When $\theta_1 < \frac{1}{4}\bar{\theta}$ or $\theta_1 > \frac{3}{4}\bar{\theta}$, we should consider the corner solution, in which $p_{2,A}^B = 0$ or $p_{2,B}^A = 0$. When $\theta_1 < \frac{1}{4}\bar{\theta}$ it follows that $p_{2,A}^B = 0$, and so Firm A should set $p_{2,A}^A$ such that $v - p_{2,A}^A - \theta_1 = v - (\bar{\theta} - \theta_1)$,

$$\begin{array}{lll} p_2^A = \bar{\theta} & p_{2,A}^A = \bar{\theta} - 2\theta_1 & p_{2,B}^A = \frac{1}{3}(3\bar{\theta} - 4\theta_1) \\ p_2^B = \bar{\theta} & p_{2,B}^B = \frac{1}{3}(3\bar{\theta} - 2\theta_1) & p_{2,A}^B = 0 \end{array}$$

If $\theta_1 > \frac{3}{4}\bar{\theta}$, the equilibrium prices are:

$$\begin{array}{lll} p_2^A = \bar{\theta} & p_{2,A}^A = \frac{1}{3}(2\theta_1 + \bar{\theta}) & p_{2,B}^A = 0 \\ p_2^B = \bar{\theta} & p_{2,B}^B = 2\theta_1 - \bar{\theta} & p_{2,A}^B = \frac{1}{3}(4\theta_1 - \bar{\theta}) \end{array}$$

Proof. See Appendix. □

Lemma 2.1 shows that if a customer chooses not to share their information in the first period, they will face uniform pricing in the second period under a *pooling* assumption. However, if they reveal information in the first period, they are confronted with behavior-based prices, including poaching prices offered by the competing firm in the second period. Lemma 2.1 demonstrates that prices are independent of λ under *pooling*. Moreover, the result $p_2^i > p_{2,i}^i > p_{2,i}^j$ holds generally, irrespective of $\theta_1 \in (0, \bar{\theta})$.[10]

When we move to the privacy decision in the first period, we need to think about the consumers' endogenous decisions about their cookies. By comparing the prices for anonymous consumers with the two prices for recognized consumers, we can show that prices for anonymous customers are always higher, such that consumers can strategically choose to share their purchasing history, in order to receive lower prices in the second period. Thus, every consumer discloses their information, which implies that the mass λ of consumers on the anonymous line is zero. Firms also form their belief accordingly, $\hat{q}(\theta) = \lambda = 0 \ \forall \theta$.

Finally, we consider price setting of firms in the first period. Similar to the second period, there are two separated lines in the first period. For the line of consumers who did not share their cookies, there is a cut-off customer $\hat{\theta}_1$, who, in the first period, is indifferent between buying from firm A at p_1^A and buying from firm B at p_1^B.[11] It is determined by $\hat{\theta}_1 = \frac{\bar{\theta}}{2} + \frac{1}{2}(p_1^B - p_1^A)$.

On the other hand, on the line of those who shared their cookies in the first period,

in order to protect the marginal customer located at θ_1. Following the same logic, when $\theta_1 > \frac{3}{4}\bar{\theta}$ it follows that $p_{2,B}^A = 0$, and thereby Firm B should set $p_{2,B}^B$ such that $v - \theta_1 = v - p_{2,B}^B - (\bar{\theta} - \theta_1)$.

[10] In the special cases of $\theta_1 = 0$ ($\theta_1 = \bar{\theta}$) we would have $p_{2,B}^A = p_{2,B}^B = p_2^A = p_2^B$ ($p_{2,A}^A = p_{2,A}^B = p_2^A = p_2^B$) under pooling.

[11] $\hat{\theta}_1$ is not influenced by the prices in the second period, since the share of those who did not disclose their information is λ, and the two firms will maximize their profits by choosing p_2^A and p_2^B which are independent of the first period.

the marginal customer, θ_1, is defined by the following equivalence,

$$v - p_1^A - \theta_1 + \left[v - p_{2,A}^B - (\bar{\theta} - \theta_1)\right] = v - p_1^B - (\bar{\theta} - \theta_1) + \left[v - p_{2,B}^A - \theta_1\right].$$

The equation represents consumers indifferent between buying from firm A at p_1^A in the first period and afterward from firm B at $p_{2,A}^B$ in the second period, and buying from firm B at p_1^B in the first period and then purchasing from firm A at $p_{2,B}^A$ in the second period. Hence, the marginal consumer is given by $\theta_1 = \frac{\bar{\theta}}{2} + \frac{3}{8}(p_1^B - p_1^A)$.

In the first period, firms maximize the overall profits, thus firm A's problem is to maximize the following term with respect to the first-period prices

$$\pi^A = \lambda p_1^A \hat{\theta}_1 + (1-\lambda) p_1^A \theta_1 + \lambda p_2^A \theta_2 + (1-\lambda)\left[p_{2,A}^A \theta_2^A + p_{2,B}^A(\theta_2^B - \theta_1)\right].$$

Similarly, firm B maximizes

$$\pi^B = \lambda p_1^B(\bar{\theta}-\hat{\theta}_1) + (1-\lambda) p_1^B(\bar{\theta}-\theta_1) + \lambda p_2^B(\bar{\theta}-\theta_2) + (1-\lambda)\left[p_{2,A}^B(\theta_1 - \theta_2^A) + p_{2,B}^B(\bar{\theta} - \theta_2^B)\right].$$

From the second-period analysis, we have obtained $p_2^A = p_2^B = \bar{\theta}$ and $\theta_2 = \frac{\bar{\theta}}{2}$. Therefore, the respective third terms in the profit functions do not affect the maximization problem. Inserting $p_{2,A}^A$, $p_{2,A}^B$, $p_{2,B}^A$, $p_{2,B}^B$, and θ_1 into the two maximization problems above we can derive the two first-order conditions,

$$\frac{\bar{\theta}}{2} + \frac{3+\lambda}{8} p_1^B - \frac{3+\lambda}{4} p_1^A - \frac{5}{16}(1-\lambda)(p_1^B - p_1^A) = 0$$

$$\frac{\bar{\theta}}{2} + \frac{3+\lambda}{8} p_1^A - \frac{3+\lambda}{4} p_1^B + \frac{5}{16}(1-\lambda)(p_1^B - p_1^A) = 0.$$

Proposition 2.1. *The optimal prices under open data environment for the competing firms in both periods are*

$$\begin{aligned}
p_1^A = p_1^B &= \frac{4}{3+\lambda}\bar{\theta} \\
p_{2,A}^A = p_{2,B}^B &= \frac{2}{3}\bar{\theta} \\
p_{2,B}^A = p_{2,A}^B &= \frac{1}{3}\bar{\theta} \\
p_2^A = p_2^B &= \bar{\theta}.
\end{aligned}$$

The marginal consumer in the first period is located at $\theta_1 = \frac{\bar{\theta}}{2}$. Consumers' strategy is to disclose data such that $q(\theta) = \lambda = 0\ \forall\theta$. Therefore, we obtain a symmetric PBE.

Proof. See Appendix. □

After deriving the pooling equilibrium in Proposition 2.1, we proceed with showing the uniqueness of this equilibrium. We include separating situations and check for potential equilibria in pure strategies. The results are summarized in the following proposition:

Proposition 2.2. *There is no separating equilibrium in pure strategies in the open data environment. The pooling equilibrium derived above is unique.*

Proof. See Appendix. □

From the results we gather that given a consumer stays anonymous in the first period, they face uniform pricing in the second period. Otherwise, they are confronted with price discrimination, which leads to the identical results as in standard behavior-based pricing with exogenous privacy where competitiors do not share information(Fudenberg and Tirole, 2000).[12] The limit cases of λ reveal an interesting insight. If $\lambda = 1$, which means that none of the consumers grants access to their cookie in the first period, this results in $p_1^A = p_1^B = \bar{\theta}$, the prices match a uniform pricing strategy. If $\lambda = 0$, which means that all consumers share their information in the first period, we get that $p_1^A = p_1^B = \frac{4}{3}\bar{\theta}$, which is a standard behavior-based pricing strategy. Therefore, for all values of $\lambda\ \in (0,1)$, p_1^A and p_1^B represent a mixture of uniform pricing and behavior-based pricing. Consumers are best off by revealing data, because they can benefit from the lower customized prices in the second period. This means consumers act strategically.[13] By revealing their information, they increase the competition between firms. Therefore, revealing cookies is a dominant strategy.

2.4.3 Exclusive Data Environment

In this section, we analyze a setting where firms only learn about cookies of customers who actually bought from them. This implies that there is exclusive data in the market, as for example consumers of firm B, might reveal their purchasing history to B, such that

[12]The results also hold if transportation costs are quadratic. Details are in Appendix.

[13]The result extends to the case where consumers are myopic. As a robustness check, we show that being strategic or myopic does not affect the consumers' decisions. See Appendix for details.

B can identify them. However, firm A does not receive the information and therefore, these consumers are anonymous to A.

The pricing strategy in the second period is distinct from the open data environment, where three different prices were set by each firm after consumers made a decision regarding their cookie choices. In comparison, in the exclusive data environment firms cannot distinguish between a competitor's customers and their own anonymous consumers. They are just one mass of non-identifiable consumers. This implies that firms cannot set a poaching price to lure consumers from each other. The pricing strategy for the second period only entails a loyalty price, $p^i_{2,i}$, and an anonymous customer price, p^i_2 for $i = A, B$. The first-period pricing is similar to the open data environment and not affected by the difference in the data environment. As before, there is a marginal consumer in the first period, $\theta_1(\mathbf{p_1})$ who is indifferent between buying from A and B. The analysis is similar to Colombo (2016). However, the essential difference is that he treats λ as a parameter, while we use it as proxy for consumers' endogenous decisions regarding their cookies.

In the exclusive data environment, the Hotelling lines cannot be separated as in the open data environment. The reason is that the anonymous price serves two functions. Firstly, it is the price for the own consumers who are not identifiable and secondly, it serves as a "poaching price" for competitors' consumers. Figure 2.3 below depicts this clearly, since p^i_2 appears on both lines. Firms want to maximize their profits by choosing prices $p^i_{2,i}$ and p^i_2 for $i = A, B$ in the second period. As before, we start by employing the *pooling* assumption. There is a share $1 - \lambda$ of consumers who choose to give access to their cookies and a share λ of consumers who hide their cookies, with λ corresponding to the probability of hiding for every consumer. For the share λ of consumers who are anonymous there is an indifferent customer located at θ_2, impartial between buying from A at price p^A_2 and B at price p^B_2. For the identifiable consumers, there is a marginal consumer in each of the companies' turfs: $\theta^{A'}_2$ is indifferent between buying from A as identifiable consumer and buying from B as anonymous customer, whereas $\theta^{B'}_2$ is the respective cut-off value on B's turf. Figure 2.3 shows this customer segmentation and price setting in a rectangle, where the two horizontal lines are again connected by λ.

From Figure 2.3 we derive the maximization problems of the companies in the second period:

$$\max_{p^A_2, p^A_{2,A}} \pi^A_2 = \max_{p^A_2, p^A_{2,A}} \lambda p^A_2 \theta_2 + (1-\lambda) p^A_{2,A} \theta^{A'}_2 + (1-\lambda) p^A_2 (\theta^{B'}_2 - \theta_1)$$
$$\max_{p^B_2, p^B_{2,B}} \pi^B_2 = \max_{p^B_2, p^B_{2,B}} \lambda p^B_2 (\bar{\theta} - \theta_2) + (1-\lambda) p^B_{2,B} (\bar{\theta} - \theta^{B'}_2) + (1-\lambda) p^B_2 (\theta_1 - \theta^{A'}_2)$$

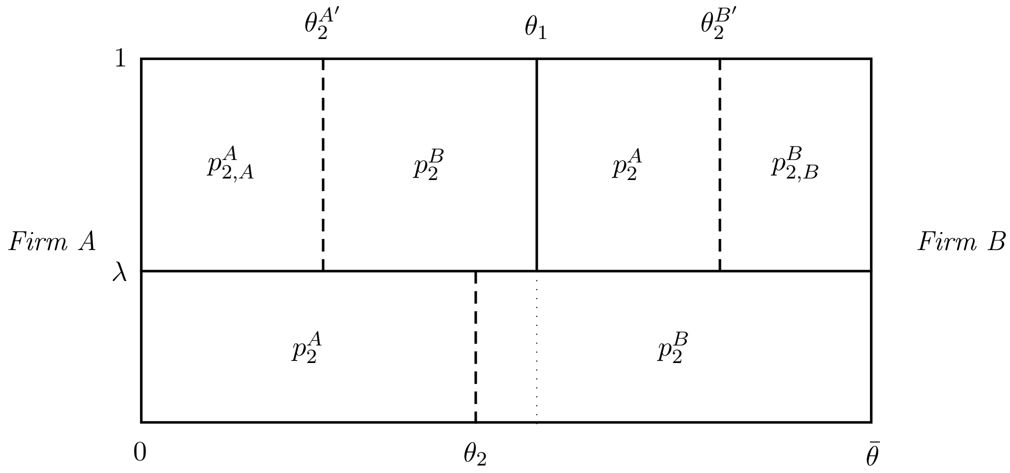


Figure 2.3: Customer segments under exclusive data

Lemma 2.2. *Solving the maximization problems, we can derive the following prices for the second period. For firm A:*

$$p_2^A(\lambda, \mathbf{p_1}) = \frac{(9 - 2\lambda + 5\lambda^2)\bar{\theta} - 4(3-\lambda)(1-\lambda)\theta_1}{3\,[4 - (1-\lambda)^2]}$$

$$p_{2,A}^A(\lambda, \mathbf{p_1}) = \frac{(3 + 10\lambda - \lambda^2)\bar{\theta} + 2(3-\lambda)(1-\lambda)\theta_1}{3\,[4 - (1-\lambda)^2]}$$

For firm B:

$$p_2^B(\lambda, \mathbf{p_1}) = \frac{(-3 + 14\lambda + \lambda^2)\bar{\theta} + 4(3-\lambda)(1-\lambda)\theta_1}{3\,[4 - (1-\lambda)^2]}$$

$$p_{2,B}^B(\lambda, \mathbf{p_1}) = \frac{(9 + 2\lambda + \lambda^2)\bar{\theta} - 2(3-\lambda)(1-\lambda)\theta_1}{3\,[4 - (1-\lambda)^2]}$$

Proof. See Appendix. □

The second-period prices in this case are not only dependent on the first-period prices, as is the case in the analysis of the open data environment, but also depend on λ as the share of buyers who choose to be anonymous.

All prices increase with λ, i.e., the more likely consumers are to hide their cookies, the higher are not only the anonymous prices but also the loyalty prices of both firms. This always holds for a line with length of one ($\bar{\theta} = 1$) and under two conditions for θ_1,

$$\theta_1 \geq \frac{1-\lambda}{3-\lambda}\bar{\theta} \qquad \theta_1 \leq \frac{2}{3-\lambda}\bar{\theta}.$$

Under these conditions, loyalty prices are larger or equal to anonymous prices, if $\theta_1 \in [\frac{1}{3}, \frac{2}{3}]$ for a normalized line. When we derive first-period prices we show that the market is separated symmetrically between the firms such that the conditions hold.

Figure 2.4 depicts loyalty and anonymous prices of firm A for $\bar{\theta} = 1$ and $\theta_1 = \frac{1}{2}$. At $\lambda = 0$ firms set symmetric prices with $p_2^A = p_2^B = \frac{1}{3}$ and $p_{2,A}^A = p_{2,B}^B = \frac{2}{3}$. The price setting corresponds to poaching and loyalty prices in the open data environment, respectively. Given $\lambda \to 1$[14] all prices converge to $\bar{\theta} = 1$, the uniform pricing strategy. There is no price discrimination in this case since there is no information available.

The graph shows that prices are convex, increasing in λ and within the range of $\lambda \in [0, 1)$, loyalty and anonymous prices do not cross. Therefore, even though the anonymous prices increase with λ, they are always below the loyalty prices. In Figure 2.4, we observe a situation that is similar to the prisoner's dilemma. Consumers face higher prices when they have a probability of $\lambda \to 1$ for anonymizing. Pricing will correspond to uniform prices. On the other hand, if consumers were to decide to hide their information with probability 0, this would lead them to a price of $\frac{2}{3}$ which is below 1. This means, if consumers can coordinate on putting zero probability on anonymizing, such that $\lambda = 0$, they would all gain. However, consumers have an incentive to deviate to stay anonymous with a positive probability, since they face an even lower anonymous price for any $\lambda > 0$. This incentive leads all consumers to anonymize.

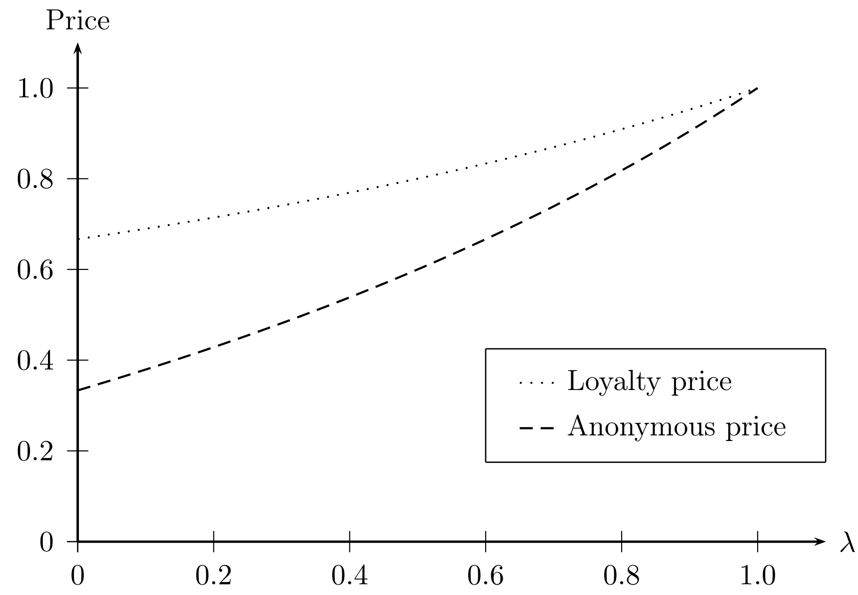


Figure 2.4: Prices of Firm A for $\bar{\theta} = 1$ and $\theta_1 = 0.5$

[14] Notice that for $\lambda = 1$, the loyalty prices are no longer contained in the maximization problems.

The spread of the two price curves is getting smaller with increasing λ. Hence, the incentive to switch to an anonymizing strategy becomes smaller.

Because consumers' best strategy is to hide their cookies with probability $\lambda = 1$, the two periods in this game are independent of each other. Therefore, in the first period the firms solve the following maximization problems:

$$
\begin{aligned}
p_1^A &= \arg\max_{p_1^A} \ \pi^A = p_1^A \theta_1 + \pi_2^A \\
p_1^B &= \arg\max_{p_1^B} \ \pi^B = p_1^A (\bar{\theta} - \theta_1) + \pi_2^B
\end{aligned}
$$

where $\theta_1 = \frac{p_1^B - p_1^A + \bar{\theta}}{2}$ for $\lambda = 1$.

Proposition 2.3. *In the exclusive data environment, final prices all coincide with the uniform pricing strategy, such that prices on the first and second period are $\bar{\theta}$. Therefore, the PBE is an equilibrium in pure strategies with $\lambda = 1$.*

Proof. See Appendix. □

Similar to the open data environment, now we prove the uniqueness of this pooling equilibrium by including the separating cases in pure strategies. The following proposition summarizes the result:

Proposition 2.4. *There is no separating equilibrium in pure strategies under exclusive data environment. The pooling equilibrium derived above is unique.*

Proof. See Appendix. □

Consumers have an incentive to choose to anonymize with highest probability, i.e. $\lambda = 1$. The results stands in contrast to the implication derived in the open data environment where all consumers give their cookies (i.e. $\lambda = 0$). In the exclusive data environment companies obtain larger profits because they do not receive information about their consumers. Therefore, the firms cannot set customized prices but have to conform to a uniform pricing strategy.

2.5 Experiment

Our experimental design consists of three components. The first and main part is a multi-stage market game, closely resembling our theoretical setup. The second part is a short and simple iterative thinking task inspired by the "Game of 21" (Dufwenberg et al., 2010, Gneezy et al., 2010). The last part is a survey on privacy concerns following Malhotra et al. (2004) to collect the IUIPC score.[15] The second and third part serve as secondary measures and controls.

Market game

Our implemented market game closely follows the theoretical setup and aims at testing our predictions concerning the buyers' privacy choices and the sellers' pricing choices under the two data environments. Subjects take the role of sellers or buyers, with roles remaining fixed for the duration of the experiment. Each market contains two sellers and six buyers and lasts for two periods. Two markets $m \in \{1, 2\}$ are simultaneously formed within one matching group, with matching groups consisting of six buyers $i \in \{1, 2, 3, 4, 5, 6\}$ and four sellers $j \in \{A, B, C, D\}$. Buyers are active in both markets, while sellers are only active in one market. This allows for a randomization of seller composition between market rounds, so that markets are independent between rounds and resemble one-shot interaction.[16]

An experimental market consists of eight adjacent locations, with sellers being located at either end and six buyers in between on distinct locations as depicted in Figure 2.5. Similar to Camacho-Cuena et al. (2005) and Barreda-Tarrazona et al. (2011) we allow sellers to choose integer prices from the interval $\in [0, 10]$. Buyers exert unit transport costs per unit of distance traveled.[17] Due to this discretization of prices and transport costs, equilibrium predictions are in pure strategies.

Buyers have an induced reservation value of 15. The utility of a buyer for a purchase is

$$U_t^i = 15 - p_t^i - \tau$$

with p_t^i describing the price of the product that buyer i chose in period t and τ as

[15]The full questionnaire is listed in the Appendix.

[16]In the comparable seller-only experiments by Brokesova et al. (2014), matching groups of four were shown to be suitable, according to Mahmood (2014) buyer involvement increases when active in multiple markets.

[17]For example, a buyer at location five has to bear transport costs of five to buy from a seller at location zero.

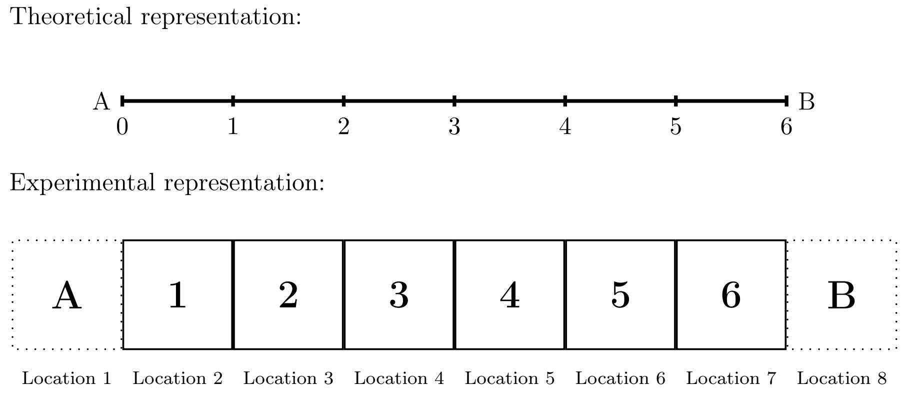


Figure 2.5: Conversion of theoretical into experimental market

transportation costs. A seller j receives the profit

$$\Pi_j = p_1^j \cdot n_1^j + \mathbf{p_2^j} \cdot \mathbf{n_2^{j'}}$$

with p_1^j corresponding to the chosen first-period price under which n_1^j is the number of buyers who bought from j. Similarly $\mathbf{p_2^j}$ is the vector of the second-period prices and $\mathbf{n_2^j}$ the vector of the number of second-period buyers who bought from j. Our two main treatment variations are (i) open data treatment and (ii) exclusive data treatment according to the two data environments in our theoretical model.

Iterative thinking task

The iterative thinking task is a variation of the "Game of 21" (Dufwenberg et al., 2010, Gneezy et al., 2010). In our version, players take turns increasing a counter that starts at 0 by increments of 1, 2 or 3. The game ends when either of two players reaches 22, where the player who picks 22 loses. Thereby, the game stays true to the original variation, where the player who picks 21 wins the game directly. Instead of using an interactive game between two subjects as intended in the original variation, we let each subject play against the computer. This is necessary in order to gather a measure on correct iterative reasoning for every subject.[18] Subjects learn that they play against the computer, without any detailed explanation on how the computer chooses. Unknown to the players, the computer avoids winning, while randomizing between the two or three

[18]If two players interact and one plays the optimal strategy, no conclusions can be drawn concerning the other player.

available options.[19]

This task serves several purposes. We suspect pricing decisions in this rather complex environment to be cognitively challenging for subjects. Heterogeneity of the subjects can lead to different observations of pricing behavior. We capture some of this heterogeneity in the capability of iterative reasoning. Likewise, buyers' privacy choices may be correlated with their ability to backward induct. Another, more pragmatic purpose is that the task generates a "mental distance" between the market game and the privacy survey.

Privacy survey

In an ensuing questionnaire, we ask participants to express their stance towards privacy and whether they are concerned about privacy breaches. The survey is based on Malhotra et al. (2004), which we use to calculate the IUIPC score. It consists of 10 statements, to which participants answer on a 7-point Likert scale from "strongly agree" to "strongly disagree".[20] Agreeing to the statement reflects a higher "privacy concern". The statements cover three broad categories. The first category covers four statements related to data *collection* by firms. The second category is related to general data *control* by data holders and consists of three statements. The last three statements cover *awareness* about data usage by firms. The statements are presented such that no two statements out of the same category succeed each other. The IUIPC score is then calculated as the equally weighted average of the average within-category scores normalized to $[0, 1]$. We use the score as a rough indication of the participants' general stance towards privacy related issues.

Predictions

We present hypotheses which are fully based on our theoretical model and suggest explicit pricing, privacy and switching patterns. In the next section, we discuss the results in light of these predictions and under consideration of the two side measures that we gather.

Hypotheses

Table 2.6 summarizes pricing, privacy and switching predictions based on our model, given the experimental parameterization. We expect that all buyers reveal their information

[19] Whenever the player is on the winning path, the computer randomizes between all three options, while only randomizing between the two options which avoid the winning path, whenever the player is not on the winning path.

[20] The full questionnaire is found in the Appendix.

in the open data treatment, while no buyer should reveal information in the exclusive data treatment.[21] When all buyers disclose their information, both setups correspond to BBPD. The opposite case, i.e., full anonymization, corresponds to the uniform pricing benchmark. While full disclosure is always better for buyers, the exclusive data treatment yields a coordination problem for buyers, since every buyer has an incentive to anonymize. As we expect full disclosure (anonymization) in the open data (exclusive data) treatment we retrieve the prediction of the anonymous price (loyalty price) as the optimal off-path response to marginal deviation from Lemma 2.3 (2.7). Just as in our theoretical model this restricts the multiplicity of equilibria.

Treatment	Open data	Exclusive data
Introduction price	8	6
Anonymous price	4	6
Loyalty price	4	6
Poaching price	2	(6)
Information disclosure	100%	0%
Share of switching	$33.\bar{3}\%$	0%

Table 2.6: Theoretical pricing, privacy and switching predictions

Aside from quantitative predictions we derive five qualitative hypotheses. The first hypothesis relates to information disclosure between treatments.

Hypothesis 1. *We expect more information disclosure in the open data treatment compared to the exclusive data treatment.*

More information in the open data treatment allows sellers to price discriminate, which in turn fosters switching.

Hypothesis 2. *We expect more switching in the open data treatment compared to the exclusive data treatment.*

The remaining three hypotheses relate to pricing patterns, with attention to comparative statics, rather than point predictions. The first relates to *second-period discounts*, which we define as any difference between the introduction price (first-period price) and loyalty, anonymous and poaching prices (second-period prices).

[21] The theory predicts full anonymization but rests on the fact that consumers are massless. Off-path deviation under the intuitive criterion leads to an indifference condition for exactly one *massless* consumer (see Lemma 2.7). Due to discretization, buyers bear a mass in the experiment. However, no matter what the number of buyers is, at most one buyer who is located centrally would disclose information.

Hypothesis 3. *We expect second-period discounts in the open data treatment, but not in the exclusive data treatment.*

For the next hypothesis we define *loyalty discounts* as a (positive) difference between anonymous price and loyalty price.

Hypothesis 4. *We expect no loyalty discounts in both treatments.*

Similarly we define *poaching discounts* as a (positive) difference between loyalty price and poaching price for the last hypothesis. Note that sellers cannot post poaching prices in the exclusive data treatment, but may "poach" by means of the anonymous price.

Hypothesis 5. *We expect poaching discounts in the open data treatment.*

In total we invited 160 students in 8 sessions of 20 each as subjects in our experiment, with 96 taking the role of buyers and 64 taking the role of sellers. On average subjects earned about 20 EUR in the 90 minutes experiment over the three parts. Most subjects were majors in economics, mathematics or industrial engineering and 36% of the subjects were female. Participants earned 2 EUR when they win against the computer in the Game of 22, which 91.25% of the participants successfully did. Lastly, participants were awarded 1 EUR for filling out the privacy survey. Sessions were conducted in the laboratory of TU Berlin and WZB in September and November 2019 with participants drawn from the ORSEE pool (Greiner, 2015). The experiment was programmed and conducted with the experiment software z-Tree (Fischbacher, 2007).

2.6 Results

Our main interest in the market game is the pricing strategies of sellers and the information disclosure by buyers. Table 2.7 shows average results of the last 10 rounds of the experiment and associated predictions. At first glance introduction prices appear to be slightly higher in the open data treatment compared to the exclusive data treatment, while both are below their predicted levels. In both treatments loyalty prices and anonymous prices are very close, which coincides with our predictions. The poaching price in the open data treatment is lower compared to the other second-period prices both within and between treatments. The share of information disclosure is larger in the open data treatment at around 2/3. Surprisingly, in more than half of the instances buyers disclosed information in the exclusive data treatment. Switching is more prevalent in the open data treatment compared to the exclusive data treatment, which is largely driven by larger prevalence of poaching in the open data treatment.

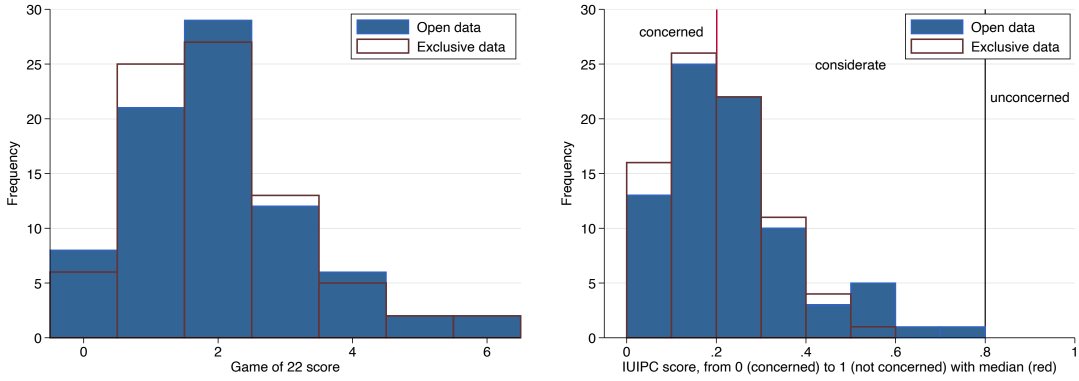
Treatment	Open data	Exclusive data
Introduction price		
Observed mean	5.55	5.54
Model prediction	8	6
Anonymous price		
Observed mean	3.85	3.79
Model prediction	4	6
Loyalty price		
Observed mean	3.94	4.07
Model prediction	4	6
Poaching price		
Observed mean	3.00	3.79
Model prediction	2	6
Information disclosure		
Observed mean	66.77%	58.54%
Model prediction	100%	0%
Share of switching		
Observed mean	23.44%	17.40%
Model prediction	$33.\bar{3}$%	0%

Table 2.7: Summary statistics for pricing, privacy and switching behavior per treatment (last 10 rounds)

In Figure 2.8 we show the distributions of the secondary measures concerning iterative thinking capabilities and privacy concern. Our findings in the Game of 22 (Figure 2.8a) are in line with Dufwenberg et al. (2010) where the majority of subjects are able to solve two steps of backward induction. In contrast to their results our subjects did not show an ability to immediately solve the game, with barely anyone solving the full six steps of induction. Overall, the results suggest that the game is suitable as a rough measure of iterative thinking capability and we cannot detect any differences between our treatments.[22]

Our findings on privacy concern are depicted in Figure 2.8b. This distribution does not show remarkable treatment differences. However, there is a tendency towards high privacy concern among our subjects. Going onwards we classify our subjects into three groups, using the median (0.2014) as an initial breaking point and 1-median ($1 - 0.2014 = 0.7986$)

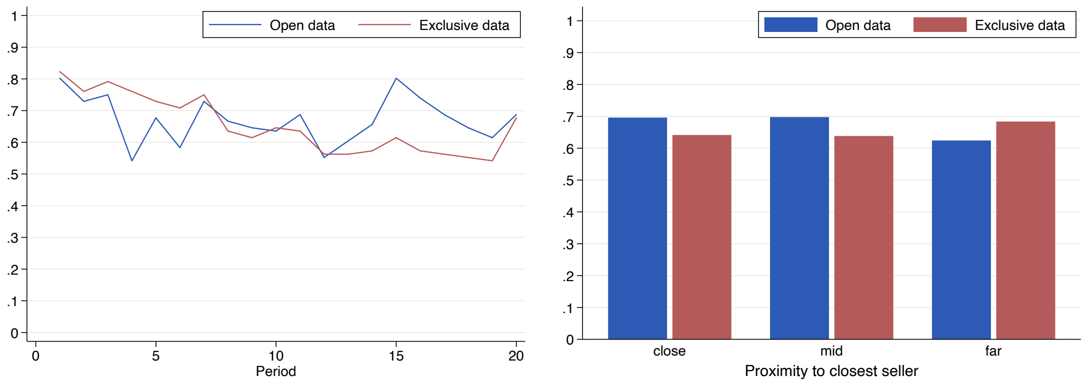
[22]We categorize subjects into three groups according to their Game of 22 score and add the information as controls in our regressions. The first group contains subjects who score below average, with either 1 or 0 steps of induction and contains 37.50% of subjects. The second group contains subjects who score the average of two steps and contains 35.00% of subjects. The last group contains all subjects who score above average that is three or higher and contains the remaining 27.50% of subjects. In Table 2.28 we show that the Game of 22 score has no impact on information disclosure.



(a) Game of 22 scores by Treatment (b) IUIPC scores by Treatment

Figure 2.8: Histograms of iterative thinking and privacy tasks

as the second breaking point. We classify a score below the median as “privacy concerned”, a score between the median and 1-median as “privacy considerate”, and a score above 1-median as “privacy unconcerned”. By nature of this classification half of our subjects fall into the first category of concerned, while surprisingly not a single subject falls into the last category of unconcerned. Thus, the remaining half of the subjects are considerates.[23] We use these measures to get a deeper understanding on what is driving the information disclosure, which on first glance appears to be unaffected by our treatment variation.



(a) Cookie choice over periods by treatment (b) Cookie choice over locations by Treatment

Figure 2.9: Cookie choice over periods and locations by treatment

Upon first inspection in Figure 2.9 we observe two things regarding the cookie choice

[23] Among the considerates we also observe a tendency leaning towards privacy concern. Moreover, irrespective of the final score all subjects expressed concern at least once within the 10-item questionnaire.

between treatments with respect to consecutive periods and different locations. With respect to periods we can see that cookie choices are initially larger and subsequently lower in the exclusive data treatment compared to the open data treatment. Albeit the similar overall sharing rate, this suggests that the adaptive processes are different between treatments. This reflects the findings of Schudy and Utikal (2017) who show that subjects are less inclined to share information, the more parties are involved in receiving information. Subjects in our experiment face a similar situation, since there are two recipients in the open data treatment and only one in the exclusive data treatment. Similarly, with respect to consumer locations, where we differentiate between close, mid and far distance to the nearest seller,[24] we find that there seems to be a subtle treatment difference. In both cases information sharing differs between the far location and the close and mid locations. However, the differences go into opposite directions between treatments. That is, in the open data treatment there is slightly less information disclosure at the far location, while in the exclusive data treatment information disclosure is slightly increased at the far location. While these effects are rather small in absolute terms we find that they are significant even under consideration of various controls.

In Table 2.10 we explore these presumptions by employing a multi-level logit model on the cookie choices of buyers, while controlling for demographics and experiment specific factors, as well as, iterative thinking capability. Specifications (1) and (2) show that there is no blunt treatment effect visible. In specifications (3) and (4) we explore the role of learning, by including a dummy variable which indicates the second half of the experiment, corresponding to rounds 11 and after.[25] There is a significant drop of information disclosure in the exclusive data treatment, while there is no change after learning in the open data treatment. Specifications (5) and (6) show the effect of location on cookie choice. We observe a significant negative effect for subjects far from the closest seller in the open data treatment. Whereas, the effect is reversed for subjects in the exclusive data treatment located at the center. The final specifications ((7) and (8)) exhibit that the effects on learning and location are independent of each other and remain unchanged under open data.

In the following we employ our classification between privacy "concerned" and "considerate" subjects. In Figure 2.11 we show the average rate of information disclosure over period by treatment and privacy concern classification. There are two major observations

[24]We consider the market to be mirrored at the half split, i.e., locations 1 and 6 are equivalent in that the proximity to the closest seller is 1 step (close). Similarly locations 2 and 5 (mid) and 3 and 4 (far) are equivalent.

[25]Results are similar when using a continuous variable indicating the round instead of the dummy for the second half.

	Dependent variable: Cookie choice $\in \{0, 1\}$							
	Treatment		Learning		Location		Learning + Location	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cookie								
Exclusive	0.036	0.607	0.442	1.028**	-0.204	0.357	0.199	0.777
	(0.278)	(0.489)	(0.295)	(0.504)	(0.299)	(0.501)	(0.315)	(0.516)
Second half			-0.046	-0.046			-0.046	-0.046
			(0.107)	(0.107)			(0.108)	(0.107)
Exclusive			-0.763***	-0.770***			-0.764***	-0.771***
× Second half			(0.156)	(0.156)			(0.156)	(0.157)
Mid					0.003	0.010	0.003	0.010
					(0.136)	(0.135)	(0.136)	(0.136)
Far					-0.326**	-0.324**	-0.326**	-0.324**
					(0.132)	(0.131)	(0.132)	(0.131)
Exclusive					0.145	0.144	0.146	0.145
× Mid					(0.192)	(0.192)	(0.194)	(0.194)
Exclusive					0.567***	0.567***	0.574***	0.573***
× Far					(0.190)	(0.190)	(0.192)	(0.192)
Market	No	Yes	No	Yes	No	Yes	No	Yes
Demographics	No	Yes	No	Yes	No	Yes	No	Yes
Privacy concern	No	Yes	No	Yes	No	Yes	No	Yes
Cognitive ability	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3840	3800	3840	3800	3840	3800	3840	3800

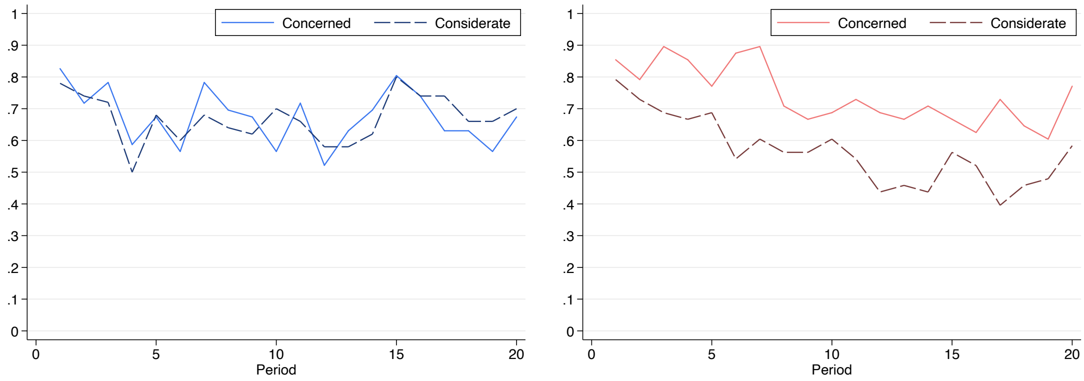
Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$
Hierarchical clustering on group and subject level.

Table 2.10: Multi-level logit on cookie choice (shortened)

here. In the open data treatment “concerned” and “considerate” buyers have a similar sharing rate. Within the exclusive data treatment “considerate” subjects always have a lower sharing rate than “concerned” subjects, which relates to the *privacy paradox*.[26] Moreover, we find that the initially large sharing rate in the exclusive data treatment is largely driven by “concerned” subjects.

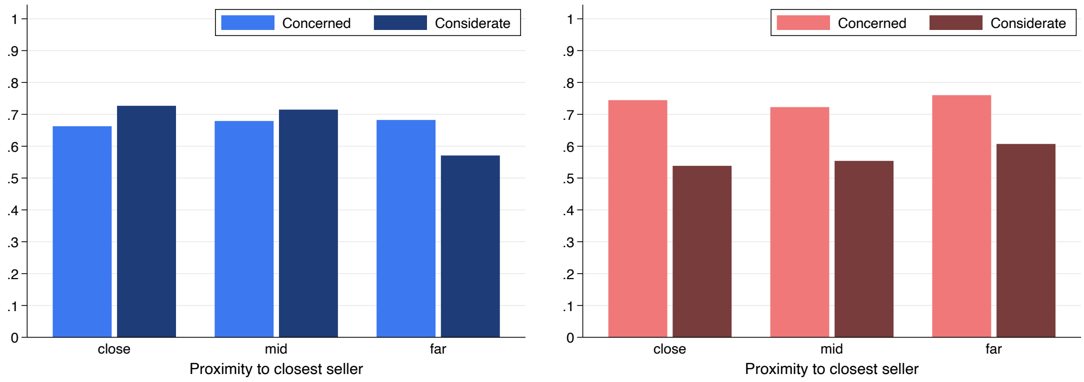
Beresford et al. (2012) and Preibusch et al. (2013) find that subjects did not act according to their stated privacy preferences when faced with a market environment. This is also reflected in Figure 2.11 where more concerned participants are actually sharing more information. However, for both privacy concerned and considerate buyers we see a drop in information sharing over periods, where especially considerate buyers in the exclusive data treatment drop below the sharing rates of the remaining three groups in the last 10 rounds.

[26] Mentioned in Acquisti et al. (2016), Dinev and Hart (2006)explain the paradox with a *privacy calculus* model, which essentially describes a mental negotiation of benefits versus concerns from disclosing information in an e-commerce setting.



(a) Open data treatment (b) Exclusive data treatment

Figure 2.11: Cookie choice over periods by treatment and privacy concern



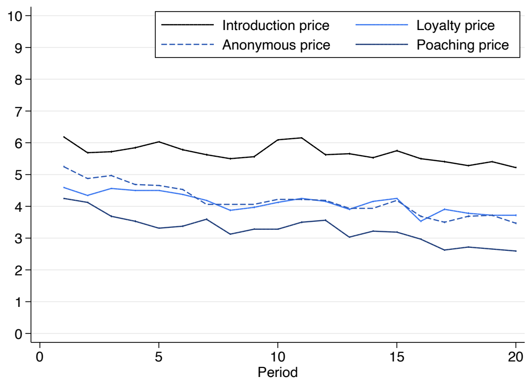
(a) Open data treatment (b) Exclusive data treatment

Figure 2.12: Cookie choice per location by treatment and privacy concern

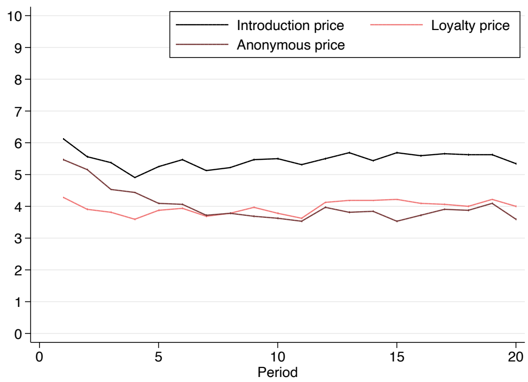
Next, we investigate whether similar discrepancies are present in the locational information disclosure between privacy types. In Figure 2.12 we show how information disclosure depends on consumers' locations under classification between privacy types. This ties in directly with our theoretical analysis which largely relies on the *pooling* assumption in the construction of equilibria. Again, we differentiate between privacy concerned and privacy considerate consumers. Figure 2.12 depicts that there is no locational preference for information disclosure in case of concerned consumers for both treatments. However, privacy considerate consumers share less information than concerned consumers in the exclusive data treatment. This is true for all locations and is considerably balanced.

The time trend did not reveal an impact of privacy concern on information disclosure in the open data treatment, but we find an impact of location in the case of considerate consumers. Considerate subjects in the open data treatment are less likely to share information at the "far" location than concerned subjects. In comparison to concerned consumers we find that considerate consumers share slightly more information in close and mid locations and less information in the far location.[27] These counteracting effects cancel out each other so that the average disclosure rate of considerate consumers is similar to the disclosure rate of concerned consumers. While we cannot explain this behavior on theoretical grounds, we can suggest that considerate consumers are more involved when it comes to disclosure of private information and both the information setting (open and exclusive data) and the individual preferences (where preferences are described by location) factor into the decision.[28] Overall, we find some evidence towards Hypothesis 1, under consideration of learning effects and differentiation of privacy concern types. That is, we observe less information sharing in the exclusive data treatment after time and mainly driven by privacy "considerate", i.e. less concerned subjects.

Our interpretation of these findings is that a high privacy concern might be the naïve "standard", which is widely adopted due to the vast media coverage and political attention privacy issues received in the past years. Those who are intrinsically involved might therefore not land in the "extreme" regions of a privacy score. Thus those who appear as "considerate" might in fact be more involved and could be regarded as the "truly concerned" subjects.



(a) Open data treatment



(b) Exclusive data treatment

Figure 2.13: Average prices per period for both treatments

[27]We cannot motivate this on behavioral grounds, but these deviations relate to Lemma 2.3 and 2.7 in that the most probable deviations are suspected in the central locations. The according response by sellers is setting loyalty price equal to anonymous price, which we explore later.

[28]All mentioned effects are stated in a regression table in the Appendix.

In Figure 2.13 we show the price paths per treatment. First we investigate price setting of sellers within treatments. We employ a fixed-effects regression, clustered on group level. We regress on respective price differences and analyze the constant remainder while considering the impact of learning, since the fixed-effects absorb any other subject-specific characteristics. In line with the descriptive summary we observe that introduction prices are larger than second-period prices. Only sellers in the open data treatment seem to employ substantial price discrimination by offering lower poaching prices for consumers who share their information and bought from the competing seller in the introduction period.

As shown in Table 2.14 we find significant differences between all second-period prices compared to first-period prices. We expected this for the open data treatment, but not for the exclusive data treatment according to Hypothesis 3. Predictions for the exclusive data treatment are rested on the fact that consumers fully anonymize. However, we observe a high rate of information sharing, which is consistent with differences between first- and second-period prices. As a second order effect this should also lead to price discrimination, which we only observe to a small extent as seen in the last column of Table 2.14.

There are significant differences between poaching price compared to loyalty and anonymous price in the open data treatment. These differences remain consistent over the course of the experiment. Initially, anonymous prices are larger than loyalty prices, while the opposite is true in the second half of the experiment as indicated by the second-half dummy for both treatments. These are the only instances of significant impacts that are reverse to the original effect. We expected no difference between loyalty and anonymous prices according to Hypothesis 4. Our observation entails only small magnitude with changing signs, which speaks in favor of the Hypothesis.[29]

The relation of information disclosure and pricing strategies is more involved. Especially in the exclusive data treatment the high rate of information sharing should have led sellers to increase their loyalty prices according to our theory. However, sellers seem reluctant to do so. We find some indication of sellers adopting poaching strategies in the exclusive data treatment towards the end of the experiment as noted before. But this seems to be limited to a selected group of sellers and was not nearly as prevalent as poaching in the open data treatment. Possibly, sellers did not understand the strategic interaction, specifically that loyal customers tend to be closer to their location. Another explanation is that sellers are driven by trust or reciprocity, such that they do not punish

[29] As we show in the Appendix, some groups of sellers begin to adopt poaching strategies in the exclusive data treatment over the course of the experiment.

	p_{intro} $-p_{anon}$	p_{intro} $-p_{loyal}$	p_{intro} $-p_{poach}$	p_{anon} $-p_{loyal}$	p_{anon} $-p_{poach}$	p_{loyal} $-p_{poach}$	p_{intro} $-p_{anon}$	p_{intro} $-p_{loyal}$	p_{anon} $-p_{loyal}$
Constant	1.266**	1.500**	2.247**	0.234**	0.981**	0.747**	1.144**	1.538**	0.394**
	(0.103)	(0.103)	(0.094)	(0.083)	(0.093)	(0.080)	(0.114)	(0.066)	(0.126)
Second half	0.434*	0.116	0.300	-0.319	-0.134	0.184	0.616*	-0.062	-0.678*
	(0.206)	(0.206)	(0.189)	(0.165)	(0.186)	(0.159)	(0.228)	(0.132)	(0.253)
Treatment		Open data			Open data			Exclusive data	
Subjects	32	32	32	32	32	32	32	32	32
Observations	640	640	640	640	640	640	640	640	640

Standard errors in parentheses
$p < .10$, * $p < .05$, ** $p < .01$
p_{intro} – Introduction price, p_{loyal} – loyalty price, p_{anon} – Anonymous price, p_{poach} – Poaching price

Table 2.14: Fixed-effects regression on price differences

loyal customers with higher prices when those were the ones who shared their information with them. In contrast, in the open data treatment poaching prices are lower compared to anonymous prices and represent a reward for sharing the information. Further, the analysis of pricing behavior of sellers sheds some light on the high rate of information disclosure in the early rounds of the exclusive data treatment, as sellers in fact offer loyalty discounts in the early rounds and only later adopt a different strategy, where they actually employ loyalty mark-ups.

Lastly, we are interested in differences in price setting behavior between treatments. We observe a larger share of sellers offering poaching discounts[30] in the open data treatment with 56.88% occurrence versus 33.75% occurrence in the exclusive data treatment. In comparison to that sellers in the exclusive data treatment were more likely to offer loyalty discounts[31] with 41.25% occurrence compared to 32.03% occurrence in the open data treatment. We measure the intensive effects on prices between treatments by random-effects regressions with group-level clustering. Since fixed-effects regression is not applicable to detect treatment differences, we use available controls in demographics, iterative thinking capability and learning effects.

In Table 2.15 we show the results. We find no significant effects on introduction, loyalty and anonymous prices. Though insignificant, the signs of all three effects correspond to

[30] We count poaching discounts as instances where loyalty prices are strictly larger than poaching prices.

[31] We conservatively count loyalty discounts as instances where loyalty prices are strictly lower than anonymous prices. Comparing loyalty prices to poaching prices yields harsher results with only 13.91% occurrences in the open data treatment, while the exclusive data treatment is unaffected since by definition *poaching price=anonymous price* in that case.

	Introduction price		Anonymous price		Loyalty price		Poaching price	
Exclusive	-0.205	-0.021	-0.173	0.167	-0.153	0.162	0.741^{*}	0.920^{*}
	(0.442)	(0.395)	(0.503)	(0.444)	(0.423)	(0.345)	(0.446)	(0.499)
Learning	No	Yes	No	Yes	No	Yes	No	Yes
Demographics	No	Yes	No	Yes	No	Yes	No	Yes
Iterative thinking	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1280	1280	1280	1280	1280	1280	1280	1280

Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$
Random-effects regression. Clustering on group level.
We use *anonymous price* $\equiv$ *poaching price* in the exclusive data treatment.

Table 2.15: Random-effects regression on treatment effects for prices

our theoretical predictions. Most notably there is a significant effect on poaching prices, indicating that sellers poach more (intensively) in the open data treatment, by roughly the same magnitude as the within-treatment analysis has revealed for the open data treatment. These results strongly confirm Hypothesis 5.

2.7 Welfare

In order to fully understand our results and their implications for policy makers, we analyze consumer and producer surplus as well as social welfare for the two informational settings. The theoretical analysis is based on the equilibria we find in Proposition 2.1 and Proposition 2.3.

The producer surplus (profit) shows that firms prefer a setting where information is not shared with a competitor.

$$\pi^*_{open} = \frac{17}{18}\bar{\theta}^2 < \pi^*_{excl.} = \bar{\theta}^2.$$

The profits are larger in the equilibrium under exclusive data. Consumers' equilibrium strategy is to anonymize, hence firms set uniform prices in both periods. Compared to the open data environment, prices are higher in the second stage of the exclusive data environment, benefitting firms. Under open data, profits are lower because in equilibrium consumers choose to accept cookies. This leads to an increase in competition between the two firms. In our model firms cannot commit to not use information about their customers. Therefore, firms prefer a setting where on equilibrium they do not receive any information about consumers.

For consumers the case is not as simple, since they receive different utilities based on their type. Utilities are determined by the information setting and the buying decisions over the two periods. It matters whether consumers are loyal to a firm over both periods or whether they were poached in the second period. The type-dependent equilibrium utilities for the different information settings are given by the following terms:

$$
\begin{aligned}
U^*_{open}(\theta) &= \begin{cases} 2(v-\bar{\theta}-\theta) & \text{for } \theta \in [0, \frac{\bar{\theta}}{3}) \\ 2v - \frac{8}{3}\bar{\theta} & \text{for } \theta \in (\frac{\bar{\theta}}{3}, \frac{2\bar{\theta}}{3}) \\ 2(v-2\bar{\theta}+\theta) & \text{for } \theta \in (\frac{2\bar{\theta}}{3}, \bar{\theta}] \end{cases} \\
U^*_{excl.}(\theta) &= \begin{cases} 2(v-\bar{\theta}-\theta) & \text{for } \theta \in [0, \frac{\bar{\theta}}{2}) \\ 2(v-2\bar{\theta}+\theta) & \text{for } \theta \in (\frac{\bar{\theta}}{2}, \bar{\theta}] \end{cases}
\end{aligned}
\tag{2.1}
$$

When comparing the utility levels of the different information settings, we find that consumers are indifferent for $\theta \in [0, \frac{\bar{\theta}}{3})$ and $\theta \in (\frac{2\bar{\theta}}{3}, \bar{\theta}]$, but obtain a higher utility for $\theta \in (\frac{\bar{\theta}}{3}, \frac{2\bar{\theta}}{3})$ from the open data environment. Consumers who are located further away from the firms can benefit from behavior-based pricing and receive a larger rent due to lower poaching prices that are available to them.

The consumer surplus shows that overall utility is larger under the open data environment.

$$
CS_{open} = 2v\bar{\theta} - \frac{22}{9}\bar{\theta}^2 \qquad CS_{excl.} = 2v\bar{\theta} - \frac{5}{2}\bar{\theta}^2
$$

$$
\Leftrightarrow CS_{open} > CS_{excl.}
$$

Consumers and firms prefer opposing information settings. Consumers' interest is to share their data with all firms on the market because firms cannot commit to not use the data. This increases competition between firms. On the other hand, firms benefit from a situation where each competitor keeps their consumers' data to themselves. The level of data available to firms drives the results.

The overall welfare level is higher under the firm-preferred exclusive data environment. The efficiency loss incurred by firms under open data is larger than the loss of consumers under exclusive data. The welfare loss under open data comes from inefficient switching, i.e., consumers that are poached do not buy from the closest firm. While consumers gain from being poached, as can be seen in the comparison of utility levels, firms lose profits (compared to the exclusive data environment) because of the lower poaching prices they set.

$$W_{open} = 2v\bar{\theta} - \frac{5}{9}\bar{\theta}^2 \qquad W_{excl.} = 2v\bar{\theta} - \frac{1}{2}\bar{\theta}^2$$
$$\Leftrightarrow W_{excl.} > W_{open}$$

The theoretical analysis shows social welfare to be lower in the open data environment, since sharing of data between firms incentivizes consumers to grant access to their data and fosters inefficient switching. However, the experimental results of the exclusive data treatment exhibit that there is a high cookie sharing rate, particularly among buyers located centric (cf. Figure 2.9). Therefore, switching does not only take place under open data, but also under exclusive data.

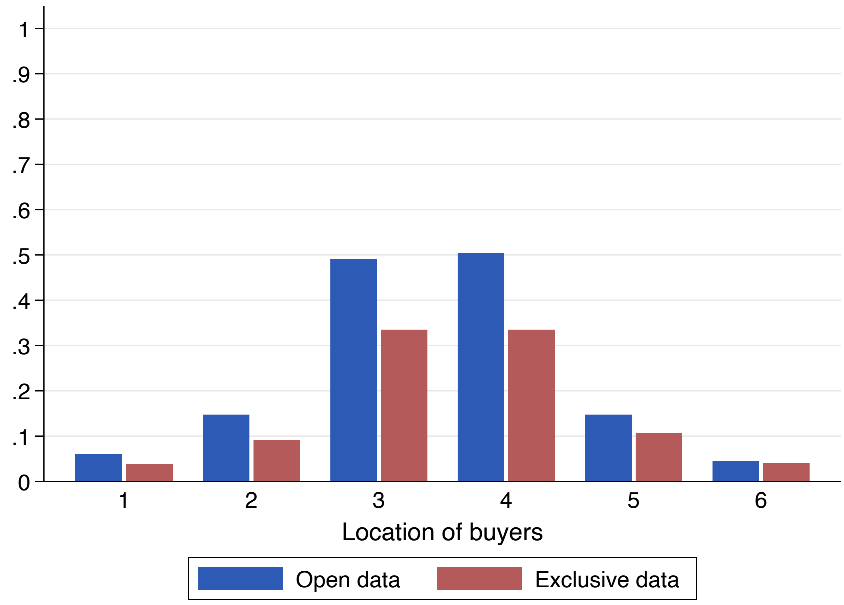


Figure 2.16: Switching by location

In Figure 2.16 we observe switching across treatments and for all locations. The further away buyers are located from the closest seller, the higher are the switching rates. These buyers benefit the most from poaching offers, which is also what we find in consumers' type-dependent utilities in theory (see Equation 2.1). Generally, rates are higher in the open data treatment compared to the exclusive data treatment. This is in line with the fact that we detect poaching efforts by sellers in the open data treatment, but not as much in the exclusive data treatment as shown in Figure 2.14. Only over time do sellers learn to use the anonymous price to poach even in the exclusive data treatment. This is one reason why we observe switching in the exclusive data treatment at all, others include price dispersion and non-rational behavior of subjects.

In order to see the effect of switching on welfare in the experiment, we analyze average transport costs and use transport costs as an inverse measure for welfare.

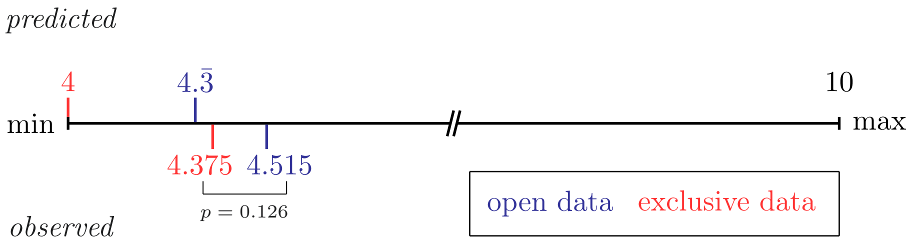


Figure 2.17: Observed and predicted average transport cost per round

Figure 2.17 depicts predicted and observed average transport cost. The left and right boundaries of the line give the average minimal and maximal transport cost per market round (i.e., two periods), respectively. This implies that the left boundary corresponds to the maximum welfare. Four is the lowest average transport cost per round if buyers purchase from the closest seller. At the same time it is the predicted value for average transport cost in the exclusive data treatment, since we expect all consumers to anonymize and not be poached by a competitor. The predicted average transport cost for the open data treatment is at $4.\bar{3}$ because we expect all subjects to share information and a portion of $\frac{1}{3}$ to switch (cf. Table 2.7). Underneath the line, we depict the observed average transport costs across treatments. The exclusive data treatment's average transport cost is at 4.375, while the open data treatment's transport cost is at 4.515. Both values are close to the maximum welfare but above the prediction of the open data treatment. Based on a random effects regression clustered on group level with time fixed effects we do not find a significant difference between the two observed average transport cost values across the treatments. Despite higher switching rates in the open data treatment compared to the exclusive data treatment, we do not observe a significant difference in social welfare. This relies on the fact that switching not only occurs inefficiently due to the poaching of customers, but also efficiently in retaining close customers who initially bought from the far seller due to market fluctuations such as price dispersion.

We find contrasting welfare results between experiment and theory, as we cannot experimentally confirm Hypothesis 2. In other words, we cannot support that welfare is higher under exclusive data. This is due to the fact that in the experiment subjects also share information in the exclusive data treatment and markets fluctuate. The theoretical welfare results hinge on the unique pooling equilibria and seem to have questionable empirical relevance.

From a policy maker's perspective, our research offers another insight. Whether to implement the mandated data sharing in specific sectors is an important issue discussed by the European Commission, and our analysis studies the effect of mandated data sharing policy when consumers have the control over their own data. The difference between open data and exclusive data environment shows that the mandated data sharing among firms will lead consumers to share more data due to the intensified competition, therefore the trade-off between privacy protection and efficiency should be attached importance when the policy maker decides on the relevant policy. Although the final decision about the mandated data sharing should be based on the specific industry, our study can at least provide some general and helpful results, on both theoretical and behavioral aspects.

2.8 Conclusion and Discussion

In this chapter, we analyze consumers' endogenous privacy decisions in a duopolistic, dynamic market where sellers employ behavior-based price discrimination. We consider two data environments, distinct in their data sharing levels. Data are contained in cookies placed by firms and contain consumers' purchasing history. In the open data environment data disclosed by consumers are fully shared between firms, whereas in the exclusive data environment data are only available to the provider. We find two unique pure-strategy equilibria. When information is open between firms all consumers fully disclose their data, which amplifies competition. Second-period prices are overall below first-period prices, particularly, firms offer poaching discounts. When information is exclusive to firms all consumers hide their data, since they are individually better off by anonymizing. Second-period prices and first-period prices correspond to uniform pricing. These results from the theoretical analysis are in line with two benchmark cases derived from earlier literature (Section 2.4.1) where the model default is exogenously full or no privacy. The welfare analysis shows that consumers prefer the open data environment, while firms prefer the exclusive data environment. Social welfare is maximal under exclusive data when consumers do not disclose data, since in absence of poaching discounts there is no inefficient switching.

We conduct a laboratory experiment closely aligned with our theoretical model where human participants act in the roles of sellers and buyers. We employ two treatments corresponding to the two data environments. We find that data sharing in the exclusive data treatment is only lower compared to the open data treatment when factoring in learning, privacy concern and location in the market. In accordance with our theory sellers price discriminate in the open data treatment by offering poaching discounts. In the exclusive data treatment sellers sparsely offer discounts to anonymous consumers even though they have access to information shared by buyers. This information is predominantly shared by privacy concerned subjects, which is discussed in the literature as the *privacy paradox*. In contrast to the theoretical prediction, social welfare in the open data treatment is not significantly lower than in the exclusive data treatment. This is mainly driven by consumers sharing data and switching providers in both treatments.

While consumers' data sharing is favorable in both data environments, there only exists an incentive to do so in the open data environment. The exclusive data environment exhibits information externalities where a collective choice of full information disclosure would lead to a better outcome for consumers, but individually consumers refrain to share information. This is also reflected in consumer's welfare which suffers under exclusive data and is higher under open data due to poaching discounts. That social welfare in theory is higher under exclusive data is solely driven by the fact that there are no inefficient switchers. However, we cannot confirm this social welfare effect in our experiment. Rather we find no significant difference in social welfare across treatments. Hence, when consumers have control over their data, policy makers should ensure an environment where information provided by consumers is shared within sectors and across firms. The European Commission is aware of this effect and proposed a legislative strategy that explores sharing data within and across sectors (European Commission, 2020).

To further explore the topic, it would be interesting to analyze a setting where consumers are in complete control of their data. Complete control entails that consumers can decide whether each firm independently receives data about their previous purchases. This way, consumers can also exclusively share their purchasing history with firms that they have not bought from. Basically, this extends the setting where both firms can place cookies (open data) by allowing consumers to choose a different option for each firm. Along this line, one can also imagine a situation of asymmetric information, i.e., a small retailer unable to collect and process consumers' data versus a large retailer accessing a wide range of personal data. It would be interesting to verify what firms and consumers' optimal strategies are when one competitor is not able to use data.

2.9 Appendix – Theoretical Part

Quadratic transportation costs

In this model, the utility for a consumer located at θ is either $v - p^i - \theta^2$ if buying from firm A, or $v - p^j - (\bar{\theta} - \theta)^2$ if buying from firm B. As in the standard model, we employ backward induction and finally find that $p_1^A = p_1^B = \frac{4}{3+\lambda}\bar{\theta}^2$, and $\theta_1 = \frac{1}{2}\bar{\theta}$, $p_{2,A}^A = p_{2,B}^B = \frac{2}{3}\bar{\theta}^2$, $p_{2,B}^A = p_{2,A}^B = \frac{1}{3}\bar{\theta}^2$. If the cost is quadratic in the standard behavior-based pricing model, prices in the first stage are $p_1^A = p_1^B = \frac{4}{3}\bar{\theta}^2$, and uniform pricing strategy is $p_1^A = p_1^B = \bar{\theta}^2$. Prices reflect quadratic transportation costs. Thus, each buyer reveals their cookies, in order to get the lower price in the second stage. λ is 0, and thus all the results in the open data environment hold with quadratic costs.

Proof of Lemma 2.1

By plugging $\theta_2^A = \frac{\bar{\theta}}{2} + \frac{p_{2,A}^B - p_{2,A}^A}{2}$ and $\theta_2^B = \frac{\bar{\theta}}{2} + \frac{p_{2,B}^B - p_{2,B}^A}{2}$ into the maximization problems, we have

$$\max_{p_{2,A}^A, p_{2,B}^A} \quad (1-\lambda)[p_{2,A}^A(\frac{\bar{\theta}}{2} + \frac{p_{2,A}^B - p_{2,A}^A}{2}) + p_{2,B}^A(\frac{\bar{\theta}}{2} + \frac{p_{2,B}^B - p_{2,B}^A}{2} - \theta_1)]$$

$$\max_{p_{2,B}^B, p_{2,A}^B} \quad (1-\lambda)[p_{2,B}^B(\bar{\theta} - \frac{\bar{\theta}}{2} - \frac{p_{2,B}^B - p_{2,B}^A}{2}) + p_{2,A}^B(\theta_1 - \frac{\bar{\theta}}{2} - \frac{p_{2,A}^B - p_{2,A}^A}{2})].$$

First-order conditions solve

$$\begin{aligned}
(1-\lambda)[\frac{\bar{\theta}}{2} + \frac{p_{2,A}^B}{2} - p_{2,A}^A] &= 0 \\
(1-\lambda)[\frac{\bar{\theta}}{2} + \frac{p_{2,B}^B}{2} - p_{2,B}^A - \theta_1] &= 0 \\
(1-\lambda)[\frac{\bar{\theta}}{2} - p_{2,B}^B + \frac{p_{2,B}^A}{2}] &= 0 \\
(1-\lambda)[\theta_1 - \frac{\bar{\theta}}{2} - p_{2,A}^B + \frac{p_{2,A}^A}{2}] &= 0
\end{aligned}$$

where we can derive the results as

$$p_2^A = \bar{\theta} \qquad p_{2,A}^A = \tfrac{1}{3}(2\theta_1 + \bar{\theta}) \qquad p_{2,B}^A = \tfrac{1}{3}(3\bar{\theta} - 4\theta_1)$$

$$p_2^B = \bar{\theta} \qquad p_{2,B}^B = \tfrac{1}{3}(3\bar{\theta} - 2\theta_1) \qquad p_{2,A}^B = \tfrac{1}{3}(4\theta_1 - \bar{\theta})$$

From these equations we observe that anonymous prices p_2^A and p_2^B are strictly positive, the same for loyalty prices $p_{2,A}^A$ and $p_{2,B}^B$. However, poaching prices $p_{2,B}^A$ and $p_{2,A}^B$ depend on θ_1 and the parameter $\bar{\theta}$. When $\frac{1}{4}\bar{\theta} \leq \theta_1 \leq \frac{3}{4}\bar{\theta}$, it is an interior solution and the equilibrium prices are just as above. When $\theta_1 < \frac{1}{4}\bar{\theta}$, it is a corner solution where $p_{2,A}^B = 0$. Firm A should set $p_{2,A}^A$ such that $v - p_{2,A}^A - \theta_1 = v - (\bar{\theta} - \theta_1)$, in order to protect the marginal customer located at θ_1. Therefore $p_{2,A}^A = \bar{\theta} - 2\theta_1$ and the other pirces are the same as in the interior solution. When $\theta_1 > \frac{3}{4}\bar{\theta}$ it follows that $p_{2,B}^A = 0$, and thereby firm B sets $p_{2,B}^B$ such that $v - \theta_1 = v - p_{2,B}^B - (\bar{\theta} - \theta_1)$. So in this case $p_{2,B}^B = 2\theta_1 - \bar{\theta}$ and the other prices do not change.

Proof of Proposition 2.1

When all consumers reveal their information, beliefs about anonymous consumers govern off-path behavior. We employ the intuitive criterion (Cho and Kreps, 1987) to support the pooling assumption.

If a single consumer individually deviates, both firms are driven to a situation of perfect competition for this single consumer, which grants the highest rent possible. Considering that both firms perfectly compete and denote $\tilde{u}(\theta)$ as the utility of a consumer of type θ who deviates:

$$\tilde{u}(\theta)^{32} = \begin{cases} v - \bar{\theta} + \theta & \text{if} \quad \theta \leq \dfrac{\bar{\theta}}{2} \\ v - \theta & \text{if} \quad \theta \geq \dfrac{\bar{\theta}}{2} \end{cases}$$

From firms' perspective, their belief about who may deviate depends on the utilities a consumer gets with and without deviation. Based on the optimal pricing strategy from Proposition 2.1 and the utility with deviation $\tilde{u}(\theta)$ derived above, we check six cases separately: when $0 \leq \theta \leq \frac{\bar{\theta}}{6}$, the utility of a consumer of type θ without deviation is $v - \frac{2}{3}\bar{\theta} - \theta$, which is larger or equal to the utility $\tilde{u}(\theta)$ if they deviate, that is $v - \bar{\theta} + \theta$. When $\frac{\bar{\theta}}{6} < \theta \leq \frac{\bar{\theta}}{3}$, the utility of a consumer of type θ without deviation is again $v - \frac{2}{3}\bar{\theta} - \theta$, which is strictly smaller than the utility if they deviate, that is $v - \bar{\theta} + \theta$. When $\frac{\bar{\theta}}{3} < \theta \leq \frac{\bar{\theta}}{2}$, the utility of a consumer of type θ without deviation becomes $v - \frac{1}{3}\bar{\theta} - (\bar{\theta} - \theta)$, which is smaller than the utility $\tilde{u}(\theta)$ if they deviate, that is $v - \bar{\theta} + \theta$. Similarly, when $\frac{\bar{\theta}}{2} < \theta \leq \frac{2}{3}\bar{\theta}$,

[32] In the perfect competition, the firm further away from the deviating consumer would set the price at zero and this consumer would be indifferent between buying from either firm.

the utility of a consumer without deviation changes to $v - \frac{1}{3}\bar{\theta} - \theta$, which is smaller than the utility with deviation, equivalent to $v - \theta$. When $\frac{2}{3}\bar{\theta} < \theta < \frac{5}{6}\bar{\theta}$, the utility of a consumer without deviation is $v - \frac{2}{3}\bar{\theta} - (\bar{\theta} - \theta)$, which is smaller than the utility with deviation $v - \theta$. Finally, when $\frac{5}{6}\bar{\theta} \leq \theta \leq \bar{\theta}$, the utility of a consumer without deviation is again $v - \frac{2}{3}\bar{\theta} - (\bar{\theta} - \theta)$, which is larger or equal to the utility with deviation $v - \theta$. Overall, we get that consumers located between $\frac{1}{6}\bar{\theta}$ and $\frac{5}{6}\bar{\theta}$ may have an incentive to deviate, thus firms form their off-path belief accordingly.

Lemma 2.3. *If firms observe a deviation of consumers' privacy choice, they believe with equal probability that it is any consumer located at $\theta \in (\frac{1}{6}\bar{\theta}, \frac{5}{6}\bar{\theta})$. The off-path price for this segment is $\frac{2}{3}\bar{\theta}$.*

Since they cannot identify the exact type of the consumer who deviates, according to the Intuitive Criterion, their belief is that the consumer with an incentive to deviate is uniformly distributed between $\frac{1}{6}\bar{\theta}$ and $\frac{5}{6}\bar{\theta}$. Therefore as a best response, they set the optimal off-path price $\frac{2}{3}\bar{\theta}$[33] if they observe a deviation. This price is equivalent to the optimal loyalty price derived in Proposition 2.1. Under these beliefs no consumer anonymizes because the total costs are not lower than under revealing information. Therefore, there is no profitable deviation for any consumer, which completes the proof.

Proof of Proposition 2.2

In this section, we prove the non-existence of a separating equilibrium in pure strategies under open data and thereby confirm the uniqueness of the pooling equilibrium derived in Proposition 2.2.

We divide all the potential scenarios into two cases: (i) when the first-period cut-off goes through a "hide" segment[34] and (ii) when the first-period cut-off goes through a "give" segment.[35] We differentiate separating equilibria according to whether the line consists of two segments or of mutiple segments. For instance, the Figure 2.18 shows the scenario of multiple segments when the first-period cut-off goes through a "hide" segment.

[33] Considering this off-path price, two firms face a continuum of consumers uniformly distributed between $\frac{1}{6}\bar{\theta}$ and $\frac{5}{6}\bar{\theta}$, thus they choose $\tilde{p}_2^A$ and $\tilde{p}_2^B$ to maximize their respective profits $\tilde{p}_2^A(\frac{\bar{\theta}}{2} + \frac{\tilde{p}_2^B - \tilde{p}_2^A}{2} - \frac{\bar{\theta}}{6})$ and $\tilde{p}_2^B[\frac{5}{6}\bar{\theta} - (\frac{\bar{\theta}}{2} + \frac{\tilde{p}_2^B - \tilde{p}_2^A}{2})]$, where we get that $\tilde{p}_2^A = \tilde{p}_2^B = \frac{2}{3}\bar{\theta}$.

[34] I.e., to the left of the cut-off all the consumers bought from firm A in the first period and to the right all bought from firm B.

[35] Please note that no assumption about symmetry is needed.

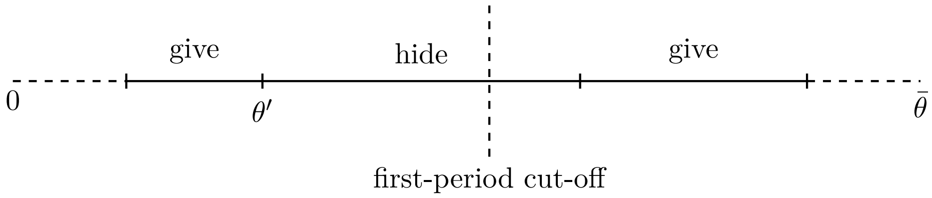


Figure 2.18: Line with multiple segments in case (i)

Definition: If not all consumers within one segment buy from the same firm, we say that there exists **Poaching Behavior** in this segment.

Lemma 2.4. *In a separating equilibrium with multiple segments, there exists no poaching behavior in any segment except for the segment that the first-period cut-off goes through, i.e., there is no poaching behavior in a lateral segment.*

Proof. We take the figure above as an example and use a proof by contradiction here. Assume Lemma 2.4 is not true and there exists poaching behavior in the left "give" segment, which means that to the left of θ' consumers buy from firm B at $p^B_{2,A}$ and to the right of θ' consumers buy from firm A at p^A_2. Since the consumer located at θ' is indifferent between revealing and hiding information, the costs of two options should be the same for them, i.e., $p^B_{2,A} + (\bar{\theta} - \theta') = p^A_2 + \theta'$. However, for those who are located to the left of θ' and buy from Firm B at $p^B_{2,A}$, they have an incentive to deviate. That is because by deviating to decline cookies, the total cost of buying from firm A would be strictly smaller than the cost before.[36] Thus, there exists a profitable deviation, which contradicts our initial assumption. The same method can be applied to a "hide" segment. This completes the proof. □

Lemma 2.4 shows that in a separating equilibrium with multiple segments there is no poaching behavior in lateral segments. Based on cut-offs between the lateral segments we can infer that $p^i_2 = p^i_{2,i}\ \forall\ i = A, B$. Now, we start to prove the non-existence of a separating equilibrium in pure strategies.

As mentioned before, we have to look at case (i) and (ii) and in each case differentiate by the number of segments (two or multiple). In other words, we need to check four possible scenarios. Let's first focus on the figure above, where there are multiple segments in case

[36] Assume that they are located at θ'' with $\theta'' < \theta'$. Since they buy from firm B at $p^B_{2,A}$, the initial costs are $p^B_{2,A} + (\bar{\theta} - \theta'')$, which is strictly larger than $p^B_{2,A} + (\bar{\theta} - \theta')$. By deviating to hide the cookies, the total costs would be $p^A_2 + \theta''$, which is strictly smaller than $p^A_2 + \theta'$. Combining together we can get that $p^A_2 + \theta'' < p^A_2 + \theta' = p^B_{2,A} + (\bar{\theta} - \theta') < p^B_{2,A} + (\bar{\theta} - \theta'')$, which shows the benefit from deviation.

(i). If we check the consumer located at θ', they are indifferent between revealing and hiding information. By Lemma 2.4, there is no poaching behavior in the "give" segments, so $p_{2,A}^A = p_2^A$. Similarly, $p_{2,B}^B = p_2^B$ also holds. However, under such circumstances both firm A and firm B have an incentive to deviate from their pricing strategy. By increasing their loyalty prices when the consumers are segmented as in the figure above, both firms could gain profit from loyal customers while keeping the profit from anonymous customers the same as before. Thus, firms have a profitable deviation and such a separating equilibrium does not exist.

Then we look at the scenario with just two segments in case (i). The Figure 2.19 below describes such a scenario:

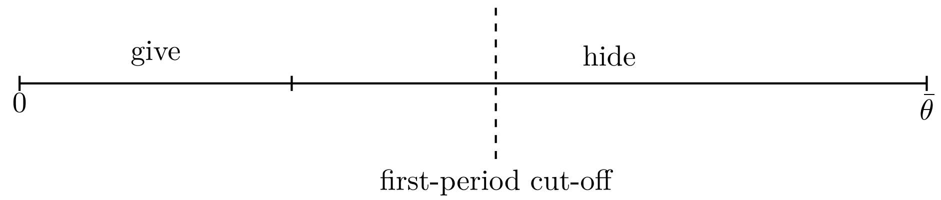


Figure 2.19: Line with two segments in case (i)

Firstly, we can show that in the "hide" segment there exists poaching behavior. Otherwise, one of the firm's anonymous prices should be 0, since both p_2^A and p_2^B are exclusively used in the "hide" segment and the firms have no reason to set a price above zero if they get no market share in this interval. If this were the case, the customers from the "give" segment would deviate to hide their cookies, since by doing so they could benefit from the zero anonymous price. Secondly, similar to Lemma 2.4 we can prove that in the "give" segment no poaching behavior exists. In other words, all consumers buy from firm A at $p_{2,A}^A$, and $p_{2,A}^A = p_2^A$. However, firm A has an incentive to raise their loyalty price, in order to obtain more from loyal customers who grant access to their cookies. Thus, this structure of separating equilibrium is not possible. Combining these two scenarios, we can conclude that in case (i) (when the first-period cut-off goes through the "hide" segment) there is no separating equilibrium in pure strategies.

In case (ii) when the first-period cut-off goes through the "give" segment, let's first look at the scenario with multiple segments along the line. In Figure 2.20 above, by Lemma 2.4, there is no poaching behavior in all "hide" segments. This means that in the left "hide" segments firm A serves all customers at a price of p_2^A and in the right "hide" segments firm B serves all at a price of p_2^B. It is similar for all "give" segments on the sides, such that $p_2^A = p_{2,A}^A$ and $p_2^B = p_{2,B}^B$. Under such circumstances both firms have an incentive to increase their anonymous prices p_2^A and p_2^B because this leads to a higher

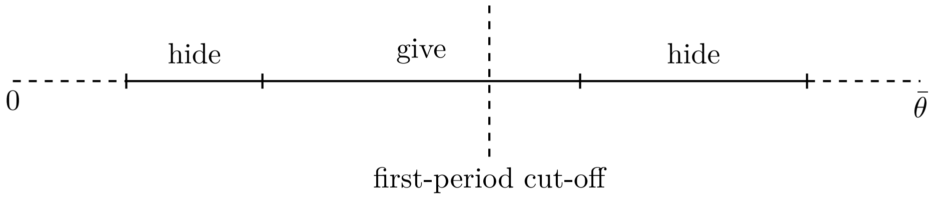


Figure 2.20: Line with multiple segments in case (ii)

profit for "hide" segments while keeping "give" segments the same as before.[37] Hence, there is no separating equilibrium in pure strategies in this scenario.

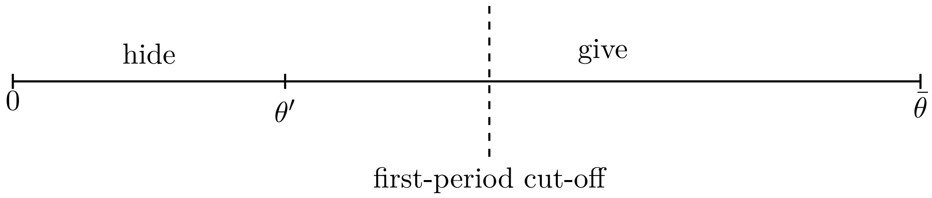


Figure 2.21: Line with two segments in case (ii)

Finally, we check the scenario with just two segments. Considering the interval between θ' and the first-period cut-off in Figure 2.21, there should exist poaching behavior. Otherwise, by the same logic mentioned before, either $p_{2,A}^{A}$ or $p_{2,A}^{B}$[38] is zero, and some outside customers deviate. Then similar to Lemma 2.4, we can easily show that no poaching behavior exists in the "hide" segment and all customers buy from firm A at p_2^A. In this condition firm A has an incentive to increase the anonymous price, in order to get more profit from the "hide" segment. To sum it up, we prove that there is no separating equilibrium in pure strategies in case (ii) when the first-period cut-off goes through the "give" segment.

All the analyses above show that there is no separating equilibrium in pure strategies in the open data environment, which completes the proof of Proposition 2.2.

Open data with myopic consumers

In the main analysis we consider consumers to be strategic. Now we want to extend our analysis to a situation in which some consumers are myopic in the first stage with regard to their purchasing decision (Baye and Sapi, 2014, Carroni et al., 2015). We

[37] This situation is similar to a Hotelling line with discontinuous demands proposed by Ackley (1942) and Shilony (1977).

[38] This interval represents those who bought from Firm A in the first stage and gave the cookies. Therefore, they are facing the loyalty price $p_{2,A}^{A}$ and poaching price $p_{2,A}^{B}$.

assume that there is a share α of myopic consumers and a share $1-\alpha$ of strategic consumers. For myopic consumers, their rationale is to choose the cheaper good in the first stage, however, they are strategic afterward, including the cookie choice and the purchasing decision in the second stage. To the contrary, strategic consumers are always forward-looking in both stages. Therefore, the difference in this setting lies in the first stage, where, among myopic consumers, marginal consumer θ_1' is just indifferent between buying from firm A at p_1^A in stage 1 and buying from firm B at p_1^B in stage 1, that is, $v - p_1^A - \theta_1' = v - p_1^B - (\bar{\theta} - \theta_1')$, leading to $\theta_1' = \frac{\bar{\theta}}{2} + \frac{p_1^B - p_1^A}{2}$. On the other hand, among strategic consumers,[39] the cut-off consumer θ_1 is indifferent between buying from firm A at p_1^A in stage 1 and then buying from firm B at $p_{2,A}^B$ in stage 2, and buying from firm B at p_1^B in stage 1 and then buying from firm A at $p_{2,B}^A$ in stage 2,[40] therefore,

$$v - p_1^A - \theta_1 + \left[v - p_{2,A}^B - (\bar{\theta} - \theta_1)\right] = v - p_1^B - (\bar{\theta} - \theta_1) + \left[v - p_{2,B}^A - \theta_1\right]$$

In order to solve this two-stage problem we apply backward induction. Starting from the second stage, again there are two separated lines for consumers who did and who did not give their cookies, respectively. No matter whether they belong to the group of myopic consumers or the group of strategic consumers, the cut-offs are the same, since even myopic consumers are strategic in the second stage. Among those who granted access to their cookies in the first stage, the two cut-offs, θ_2^A and θ_2^B, are equivalent to $\frac{\bar{\theta}}{2} + \frac{p_{2,A}^B - p_{2,A}^A}{2}$ and $\frac{\bar{\theta}}{2} + \frac{p_{2,B}^B - p_{2,B}^A}{2}$, respectively.[41] Moreover, for those who did not give their cookies in the first stage, as we discussed before, they will face uniform pricing in the second stage, with $p_2^A = p_2^B = \bar{\theta}$ and $\theta_2 = \frac{\bar{\theta}}{2}$.

Therefore, the competitors maximize their profits from the line with mass $1-\lambda$ as follows

$$\max_{p_{2,A}^A, p_{2,B}^A} \quad \alpha(1-\lambda)\left[p_{2,A}^A \theta_2^A + p_{2,B}^A(\theta_2^B - \theta_1')\right] + (1-\alpha)(1-\lambda)\left[p_{2,A}^A \theta_2^A + p_{2,B}^A(\theta_2^B - \theta_1)\right]$$

$$\max_{p_{2,B}^B, p_{2,A}^B} \quad \alpha(1-\lambda)\left[p_{2,B}^B(\bar{\theta} - \theta_2^B) + p_{2,A}^B(\theta_1' - \theta_2^A)\right] + (1-\alpha)(1-\lambda)\left[p_{2,B}^B(\bar{\theta} - \theta_2^B) + p_{2,A}^B(\theta_1 - \theta_2^A)\right]$$

Lemma 2.5. *Combining these two optimization problems and deriving the first order*

[39] To make it more precise, strategic consumers mean those who are forward-looking and reveal their data in the first stage.

[40] This indifference condition is the same as under open data with strategic consumers.

[41] The method to derive these cut-offs are identical to the open data environment with strategic consumers.

conditions, we obtain the following prices in the second stage

$$p^A_{2,A} = \frac{\bar{\theta}}{3} + \frac{2}{3}\theta_1 + \frac{2}{3}\alpha(\theta_1' - \theta_1) \qquad p^A_{2,B} = \bar{\theta} - \frac{4}{3}\theta_1 + \frac{4}{3}\alpha(\theta_1 - \theta_1')$$
$$p^B_{2,B} = \bar{\theta} - \frac{2}{3}\theta_1 + \frac{2}{3}\alpha(\theta_1 - \theta_1') \qquad p^B_{2,A} = -\frac{\bar{\theta}}{3} + \frac{4}{3}\theta_1 + \frac{4}{3}\alpha(\theta_1' - \theta_1)$$

Note that on the line with consumer mass λ nothing changes and therefore the prices correspond to uniform pricing.

On the first stage, the cut-offs are different among the myopic consumers and strategic consumers, and also depend on whether they decline cookies or not. Therefore, there are four groups of different consumers. Among the mass of λ consumers who do not share their information, a mass of $\alpha\lambda$ are myopic and a mass of $(1-\alpha)\lambda$ are strategic. However, no matter whether they are myopic or strategic, the cut-offs they face are the same, that is $\theta_1' = \frac{\bar{\theta}}{2} + \frac{p_1^B - p_1^A}{2}$.[42] Similarly, among the mass of $1-\lambda$ consumers, there are $\alpha(1-\lambda)$ myopic consumers facing the cut-off of θ_1', while a mass of $(1-\alpha)(1-\lambda)$ are strategic consumers with the cut-off of θ_1.

Combining these indifference conditions and the results from Lemma 2.5, we obtain $\theta_1' = \frac{\bar{\theta}}{2} + \frac{p_1^B - p_1^A}{2}$ and $\theta_1 = \frac{\bar{\theta}}{2} - \frac{4\alpha-3}{8(1-\alpha)}(p_1^B - p_1^A)$. Maximizing the overall profits in the first period with respect to the first-stage prices, the two firms have the resulting objective functions

$$\begin{aligned}
\pi^A &= \alpha\big[\lambda p_1^A\theta_1' + (1-\lambda)p_1^A\theta_1' + \lambda p_2^A\theta_2 + (1-\lambda)p^A_{2,A}\theta_2^A + (1-\lambda)p^A_{2,B}(\theta_2^B - \theta_1')\big] \\
&\quad + (1-\alpha)\big[\lambda p_1^A\theta_1' + (1-\lambda)p_1^A\theta_1 + \lambda p_2^A\theta_2 + (1-\lambda)p^A_{2,A}\theta_2^A + (1-\lambda)p^A_{2,B}(\theta_2^B - \theta_1)\big] \\
\pi^B &= \alpha\big[\lambda p_1^B(\bar{\theta} - \theta_1') + (1-\lambda)p_1^B(\bar{\theta} - \theta_1') + \lambda p_2^B(\bar{\theta} - \theta_2) + (1-\lambda)p^B_{2,B}(\bar{\theta} - \theta_2^B) \\
&\quad + (1-\lambda)p^B_{2,A}(\theta_1' - \theta_2^A)\big] + (1-\alpha)\big[\lambda p_1^B(\bar{\theta} - \theta_1') + (1-\lambda)p_1^B(\bar{\theta} - \theta_1) + \lambda p_2^B(\bar{\theta} - \theta_2) \\
&\quad + (1-\lambda)p^B_{2,B}(\bar{\theta} - \theta_2^B) + (1-\lambda)p^B_{2,A}(\theta_1 - \theta_2^A)\big]
\end{aligned}$$

Lemma 2.6. *Substituting the respective prices into the system of equations given by the first-order conditions, we derive the final results for the first- and second-stage prices:*

$$\begin{array}{ll}
p_1^A = p_1^B = \frac{4}{3+\lambda}\bar{\theta} & p^A_{2,A} = p^B_{2,B} = \frac{2}{3}\bar{\theta} \\
p^A_{2,B} = p^B_{2,A} = \frac{1}{3}\bar{\theta} & p_2^A = p_2^B = \bar{\theta}
\end{array}$$

Everyone chooses to give cookies, therefore the optimal λ is 1 and the resulting prices are

[42] θ_1' will not be influenced by the prices in the second stage, which is similar to the open data environment with strategic consumers.

identical to the open data environment with strategic consumers.

The result above is a robustness check, showing that being strategic or myopic in the first period purchase does not affect any decisions. Consumers choose to grant firms' access to their cookies, in order to benefit from competition; while firms use standard behavior-based price discrimination to maximize their profits. Moreover, the strategic cookie choice is sufficient to yield identical results including first period prices. This is not the case in standard BBPD models without the cookie stage.

Proof of Lemma 2.2

When maximizing the profit functions of the second stage, we get the following expressions for the first-order conditions:

$$\begin{aligned}
\frac{\partial \pi_2^A}{\partial p_2^A} &= \frac{\lambda}{2}(p_2^B - 2p_2^A + \bar{\theta}) + \frac{(1-\lambda)}{2}(p_{2,B}^B - 2p_2^A + \bar{\theta}) - (1-\lambda)\bar{\theta} = 0 \\
\frac{\partial \pi_2^A}{\partial p_{2,A}^A} &= \frac{(1-\lambda)}{2}(p_2^B - 2p_{2,A}^A + \bar{\theta}) = 0 \\
\frac{\partial \pi_2^B}{\partial p_2^B} &= \frac{\lambda}{2}(-2p_2^B + p_2^A + \bar{\theta}) + \frac{(1-\lambda)}{2}(-2p_2^B + p_{2,A}^A - \bar{\theta} + 2\theta_1) = 0 \\
\frac{\partial \pi_2^B}{\partial p_{2,B}^B} &= \frac{(1-\lambda)}{2}(-2p_{2,B}^B + p_2^A + \bar{\theta}) = 0
\end{aligned}$$

This gives a system of equations, where prices are dependent on each other and need to be substituted into each other in order to receive the final set of prices of the second period that are only depending on λ and $\mathbf{p_1}$.

$$\begin{aligned}
p_{2,A}^A(p_2^B) &= \frac{\bar{\theta} + p_2^B}{2} \\
p_2^A(p_2^B) &= \frac{(3-\lambda)\bar{\theta} + 2\lambda \cdot p_2^B - 4(1-\lambda)\theta_1}{3+\lambda} \\
p_{2,B}^B(p_2^A) &= \frac{\bar{\theta} + p_2^A}{2} \\
p_2^B(p_2^A) &= \frac{-(1-3\lambda)\bar{\theta} + 2\lambda \cdot p_2^A + 4(1-\lambda)\theta_1}{3+\lambda}
\end{aligned}$$

Please note that unlike Lemma 2.1 we do not need to consider the corner solution here. In the open data environment, the poaching price from one firm may be zero, but in the exclusive data environment, firms cannot poach and use the anonymous price instead. For any firm i, setting the price p_2^i at zero is a weakly dominated strategy since its marginal cost is just zero. However, when the information is exclusive, the anonymous price is also applied to those who hide their cookies. Considering the firm i again, if the anonymous price from another firm p_2^j is not zero, they always have an incentive to set the price above zero in order to get some profits from those who hide their cookies, which will make them strictly better off than choosing the corner solution. Thus, we do not consider the corner solution in the exclusive data environment. All the equations above can easily derive the results in Lemma 2.2.

Proof of Proposition 2.3

In the case where all consumers hide their information, the counterfactual of consumers who disclose information is governed by off-path beliefs. Suppose, without loss of generality, a single (atom-less) consumer who bought from A in period 1 deviates by disclosing information. The price setting of firm B remains unchanged, since B cannot target the deviating consumer and the impact on the price is negligible. This consumer who identifies towards A can only be located on $[0, \theta_1]$ and will receive a price from firm A to make him indifferent between buying from firm A and firm B. Thus, the utility $\tilde{u}(\theta)$ of a consumer of type θ who deviates is:

$$\tilde{u}(\theta)^{43} = \begin{cases} v - 2\bar{\theta} + \theta & \text{if} \quad \theta \leq \dfrac{\bar{\theta}}{2} \\ \\ v - \bar{\theta} - \theta & \text{if} \quad \theta \geq \dfrac{\bar{\theta}}{2} \end{cases}$$

From the firm's side, their belief about who may deviate depends on the utilities that the consumer get with and without deviating. By Proposition 2.3 and the utility with deviation $\tilde{u}(\theta)$ derived above, we check different scenarios separately: when $0 \leq \theta < \frac{\bar{\theta}}{2}$, the utility of a consumer of type θ without deviation is $v - \bar{\theta} - \theta$, which is strictly larger than the utility $\tilde{u}(\theta)$ if they deviate, that is $v - 2\bar{\theta} + \theta$. When $\frac{\bar{\theta}}{2} < \theta \leq \bar{\theta}$, the utility of a consumer of type θ without deviation is $v - \bar{\theta} - (\bar{\theta} - \theta)$, which is strictly larger than the utility $\tilde{u}(\theta)$ if they deviate, that is $v - \bar{\theta} + \theta$. Only when $\theta = \frac{\bar{\theta}}{2}$, the utility of a consumer of type θ does not change with or without the deviation.

Lemma 2.7. *If firms observe any deviation from consumers, they form the off-path belief that it is the consumer located at $\frac{\bar{\theta}}{2}$ and set the off-path price $\bar{\theta}$ as a best response.*

Since only the consumer in the center of the line gets the same utility from deviating,

[43] Please note the firm that the deviating consumer did not buy from in the first stage sets the price at $\bar{\theta}$ and this consumer is indifferent between buying from either firm.

firms' best response is to set the uniform price $\bar{\theta}$. However, consumers do not benefit from the deviation, since the utility does not change. Therefore, the proof is complete.

Proof of Proposition 2.4

In this section, we look at the exclusive data environment, where similar arguments compared to the proof of Proposition 2.2 are applied to prove that there is no separating equilibrium in pure strategies. To do so we first expand Lemma 2.4 to the case of exclusive data.

Lemma 2.8. *If there exists a separating equilibrium for a line with multiple segments under exclusive data, there is no poaching behavior in the lateral segments.*

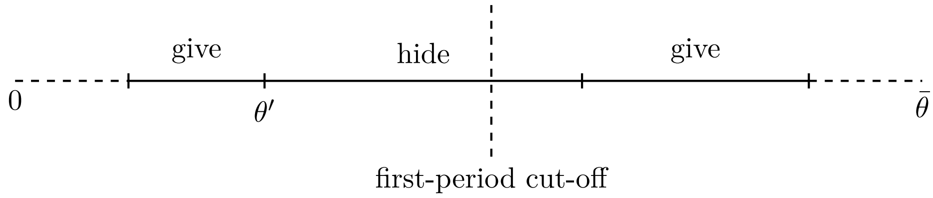


Figure 2.22: Line with multiple segments in case (i)

Proof. Let's take Figure 2.22 as an example where the first-period cut-off divides a "hide" segment. Assume towards a contradiction that there exists poaching behavior in the left segment. This means that consumers to the left of θ' buy from firm B at p_2^B and to the right of θ' buy from firm A at p_2^A. The consumer located at θ' is indifferent between accepting and declining cookies. The cost of each option should be the same for this indifferent consumer, i.e., $p_2^B + (\bar{\theta} - \theta') = p_2^A + \theta'$. However, consumers located to the left of θ' who buy from firm B at p_2^B, have an incentive to deviate. By deviating to hide cookies, the total cost of buying from firm A would be strictly smaller than the cost before.[44] The same method can be applied to the case when the first-period cut-off divides the "give" segment, which together shows that if there is poaching behavior in the lateral segments, consumers have an incentive to deviate, such that a separating equilibrium in pure strategies cannot exist. □

[44] Assume that they are located at θ'' with $\theta'' < \theta'$. Since they buy from Firm B at p_2^B, the initial costs are $p_2^B + (\bar{\theta} - \theta'')$, which is strictly larger than $p_2^B + (\bar{\theta} - \theta')$. By deviating to hide their data, the total costs would be $p_2^A + \theta''$, which is strictly smaller than $p_2^A + \theta'$. Combining together we can get that $p_2^A + \theta'' < p_2^A + \theta' = p_2^B + (\bar{\theta} - \theta') < p_2^B + (\bar{\theta} - \theta'')$, which shows the benefit from deviation.

From Lemma 2.8 we can generally infer that in a separating equilibrium firms give up their option to price discriminate since the cut-offs between lateral segments the following must hold: $p_2^i = p_{2,i}^i$, $\forall i = A, B$. Based on Lemma 2.8, we show the non-existence of a separating equilibrium in pure strategies under exclusive data.

Similar to the previous proof of Proposition 2.2, we distinguish between two cases: (i) when the first-period cut-off divides a "hide" segment, and (ii) when the first-period cut-off divides a "give" segment. Combining with the number of the segments along the line, we need to, again, check four possible scenarios separately.

Let's first consider case (i) with multiple segments. As mentioned in Figure 2.22, we assume towards a contradiction that there is a separating equilibrium with pure strategies. In order for such a separating equilibrium to exist Lemma 2.8 must hold and poaching behavior in lateral segments is excluded, which means that firms cannot poach with their anonymous prices in "give" segments. Again, this implies that firms give up the option to price discriminate in pure-strategy separating equilibria, which is directly shown from $p_2^i = p_{2,i}^i$. Yet, it is obvious that firms have an incentive to price discriminate on consumers who share their cookies. By increasing their loyalty prices when consumers are segmented as in the figure above, firms gain by giving up less rent to the consumers. This is a direct contradiction to the existence of a possible separating equilibrium.

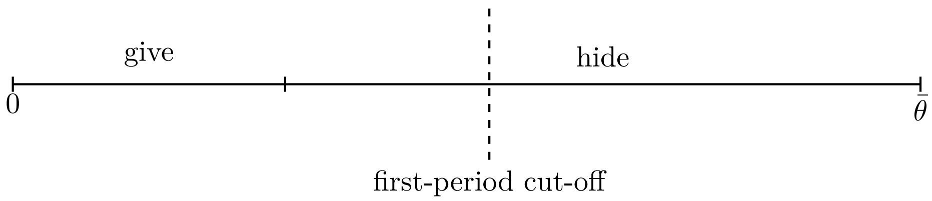


Figure 2.23: Line with two segments in case (i)

Then we look at the scenario with just two segments in case (i). If there exists such a separating equilibrium as in Figure 2.23, from Lemma 2.8 we find no poaching behavior in the "give" segment. Moreover, based on the customer indifferent between hiding and accepting cookies, we have $p_{2,A}^A = p_2^A$. Please note that under such circumstances the firms again give up price discrimination. However, firm A has an incentive to raise the loyalty price $p_{2,A}^A$. By doing so they get more profit from loyal customers and not affect the profit from anonymous customers. Therefore, this structure is not possible and we can conclude that there is no separating equilibrium in pure strategies in case (i).

In case (ii) when the first-period cut-off divides a "give" segment, let's first look at the scenario with multiple segments along the line. In Figure 2.24, by Lemma 2.8 we know that there is no poaching behavior in the lateral segments. This means that $p_{2,A}^A = p_2^A$

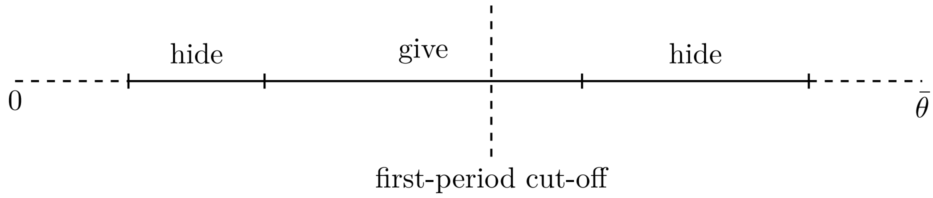


Figure 2.24: Line with multiple segments in case (ii)

and $p^B_{2,B} = p^B_2$, and firms give up their option to price discriminate in the lateral segments. Now, we focus on the central "give" segment. To the left of the first-period cut-off, all consumers face $p^A_{2,A}$ from firm A and p^B_2 from firm B. Similarly, to the right of this cut-off, all consumers choose between p^A_2 from firm A and $p^B_{2,B}$ from firm B. Given the fact that $p^A_{2,A} = p^A_2$ and $p^B_{2,B} = p^B_2$, the second-period cut-offs in these two intervals coincide, which means that $\theta^A_2 = \theta^B_2$. Considering the location of θ^A_2 and θ^B_2, there are three possibilities: to the left of the first-period cut-off, to the right of the first-period cut-off, and coinciding with the first-period cut-off.[45] If θ^A_2 and θ^B_2 are to the left of the first-period cut-off, no consumers located to the right of the first-period cut-off buy from firm A at p^A_2. However, in such a condition, firm A has an incentive to raise p^A_2 in order to get more profit. Thus, we can rule out this possibility. Similarly, if θ^A_2 and θ^B_2 are to the right of the first-period cut-off, no consumers located to the left of the first-period cut-off will buy from firm B at p^B_2 and firm B would like to increase their anonymous price. Therefore, this possibility is also excluded. Finally, if θ^A_2 and θ^B_2 coincide with the first-period cut-off, no customers in the central "give" segment buy from firm A at p^A_2 or from firm B at p^B_2. Under such circumstances, both firm A and firm B have an incentive to raise their anonymous prices and they benefit from this deviation. Overall, we have shown that no separating equilibrium exists in this scenario.

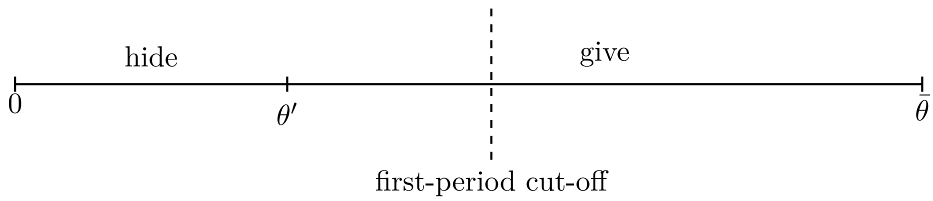


Figure 2.25: Line with two segments in case (ii)

Finally, we check the scenario with just two segments when the first-period cut-off divides the "give" segment. Firstly, in the "give" segment, there should be some consumers

[45] Please note that θ^A_2 and θ^B_2 do not need to be within the central "give" segment. All results hold even if they are not within the central "give" segment.

buying from firm A at $p^A_{2,A}$ and some buying from firm B at $p^B_{2,B}$. Otherwise, since $p^A_{2,A}$ and $p^B_{2,B}$ are exclusively set in this segment, either $p^A_{2,A}$ or $p^B_{2,B}$ should be zero and some outside consumers will deviate to this interval. Then, similar to Lemma 2.8, we can easily show that there is no poaching behavior in the "hide" segment and $p^A_2 = p^A_{2,A}$. Under such circumstances, there are two groups of consumers buying from firm A at p^A_2 on this line: those who choose to decline cookies and those who accept cookies and buy from firm B in the first stage. Apparently, due to the higher transportation cost, the second group of consumers has an incentive to deviate. They would choose to hide cookies in the first stage, and the structure of this separating equilibrium collapses accordingly. As a summary, we can conclude that no separating equilibrium in pure strategies exists in case (ii). This completes the proof.

2.10 Appendix – Experimental Part

Instructions for the experiment

Market game - Exclusive data [Open data][46]

A market

Participants take the role of buyers or sellers and are active in a market with eight locations. Two sellers sell the same good and are located on either end of the market. Six buyers are located between the two sellers according to the following graphical depiction:

Seller	**Buyer**	**Buyer**	**Buyer**	**Buyer**	**Buyer**	**Buyer**	**Seller**
Location 1	Location 2	Location 3	Location 4	Location 5	Location 6	Location 7	Location 8

Buyers buy exactly one good in each of the two periods. Sellers choose prices p at the beginning of each period. Prices must be integers between 0 and 10. Buyers pay the price of a good and transport costs t according to their distance to the respective seller. Buyers pay transport costs of one unit per field and have to move to the sellers' location. Buyers receive earnings according to the following earnings function:

$$Earnings = 15 - p - t$$

[46] Here you find translated short versions of the instrutions for the experiment. Original instructions are in German and can be made available upon request. Note that transportation costs in the instructions are denoted by t which corresponds to θ in the main analysis.

At the beginning of the first period sellers choose an *introduction price*. Buyers choose one seller and decide whether to allow *cookies*. At the beginning of the second period sellers choose three prices: a *loyalty price*[, a *poaching price*] and a *new customer price*. The profit of sellers in a market corresponds to the sold number of goods multiplied with their respective prices according to the following profit function:

$$\mathit{Profit} = p \cdot n$$

The following table depicts which buyer sees which price of the two sellers in the second period, according to their initial purchasing decision and cookie choice.

Chosen seller in first period	Allow use of cookies	Price of seller 1	Price of seller 2
Seller 1	allow	Loyalty price	New customer price
Seller 1	don't allow	New customer price	New customer price
Seller 2	allow	New customer price	Loyalty price
Seller 2	don't allow	New customer price	New customer price

[Differences in the open data treatment <u>underlined</u>.]

Chosen seller in first period	Allow use of cookies	Price of seller 1	Price of seller 2
Seller 1	allow	Loyalty price	<u>Poaching price</u>
Seller 1	don't allow	New customer price	New customer price
Seller 2	allow	<u>Poaching price</u>	Loyalty price
Seller 2	don't allow	New customer price	New customer price

Procedure

At the beginning of the experiment each participant is assigned a role, which remains fixed for the remainder of the experiment of 20 rounds in total. In each round there are two markets with two sellers each. Six buyers are active in **both** markets, while sellers are active in **one** of the markets. Within one round locations of buyers and sellers are fixed. Each round buyers are assigned random new locations in both markets. Sellers are randomly assigned to one market with a random location at either end of the market in each round.

Privacy concern survey - IUIPC score[47]

All statements are rated by the subjects on a seven-point scale from “strongly agree” to “strongly disagree”. The first three statements relate to control issues, statements four to six relate to awareness and the remaining four statemtents relate to collection issues.

1) Consumer online privacy is really a matter of consumers’ right to exercise control and autonomy over decisions about how their information is collected, used, and shared.
2) Consumer control of personal information lies at the heart of consumer privacy.
3) I believe that online privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction.
4) Companies seeking information online should disclose the way the data are collected, processed, and used.
5) A good consumer online privacy policy should have a clear and conspicuous disclosure.
6) It is very important to me that I am aware and knowledgeable about how my personal information will be used.
7) It usually bothers me when online companies ask me for personal information.
8) When online companies ask me for personal information, I sometimes think twice before providing it.
9) It bothers me to give personal information to so many online companies.
10) I’m concerned that online companies are collecting too much personal information about me.

Iterative thinking task - The game of 22[48]

The rules of the game are as follows: This is a two-player game in which players increase a counter. This counter starts at 0 and ends at 22 and must be moved each turn by 1, 2 or 3 steps, with players acting sequentially. You will play this game against the computer and you are the first to move. The player who reaches 22 loses. If the computer loses the game, you will earn 2€, while you will earn 0€ if you lose.

0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**

Figure 2.26: Representation of the Game of 22

[47] Original questions of Malhotra et al. (2004) were translated to German.
[48] Instructions are originally in German and presented on screen.

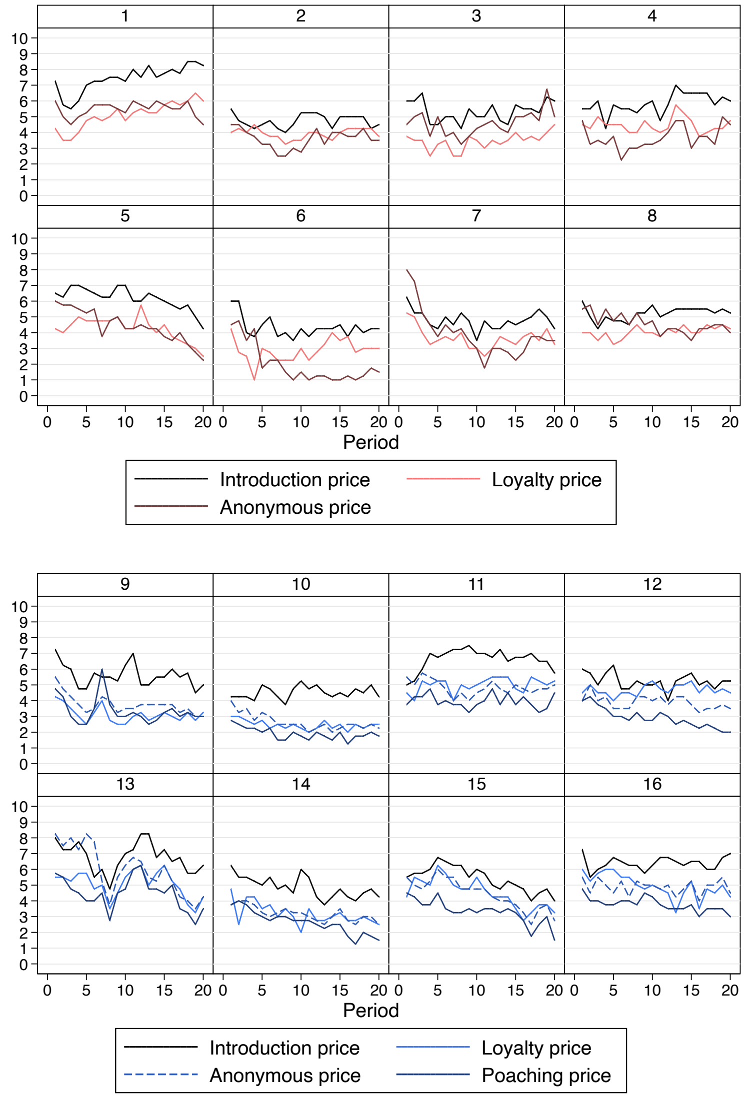


Figure 2.27: Average prices per period for exclusive data (1-8) and open data (9-16)

	Dependent variable: Cookie choice $\in \{0,1\}$							
	Treatment		Learning		Location		Learning + Location	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cookie								
Exclusive	0.036	0.607	0.442	1.028**	-0.204	0.357	0.199	0.777
	(0.278)	(0.489)	(0.295)	(0.504)	(0.299)	(0.501)	(0.315)	(0.516)
Second half			-0.046	-0.046			-0.046	-0.046
			(0.107)	(0.107)			(0.108)	(0.107)
Exclusive			-0.763***	-0.770***			-0.764***	-0.771***
× Second half			(0.156)	(0.156)			(0.156)	(0.157)
Mid					0.003	0.010	0.003	0.010
					(0.136)	(0.135)	(0.136)	(0.136)
Far					-0.326**	-0.324**	-0.326**	-0.324**
					(0.132)	(0.131)	(0.132)	(0.131)
Exclusive					0.145	0.144	0.146	0.145
× Mid					(0.192)	(0.192)	(0.194)	(0.194)
Exclusive					0.567***	0.567***	0.574***	0.573***
× Far					(0.190)	(0.190)	(0.192)	(0.192)
Considerate		-0.167		-0.170		-0.177		-0.180
		(0.351)		(0.357)		(0.351)		(0.357)
Exclusive		-1.224**		-1.267**		-1.218**		-1.262**
× Considerate		(0.492)		(0.500)		(0.491)		(0.500)
G22 score = 2		0.579		0.580		0.571		0.572
		(0.379)		(0.385)		(0.379)		(0.385)
G22 score > 2		0.056		0.056		0.056		0.056
		(0.465)		(0.472)		(0.465)		(0.472)
Exclusive		0.446		0.480		0.464		0.499
× G22 score = 2		(0.569)		(0.579)		(0.569)		(0.579)
Exclusive		-0.547		-0.559		-0.546		-0.559
× G22 score > 2		(0.620)		(0.630)		(0.620)		(0.630)
Age	No	Yes	No	Yes	No	Yes	No	Yes
Study	No	Yes	No	Yes	No	Yes	No	Yes
Gender	No	Yes	No	Yes	No	Yes	No	Yes
Market	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3840	3800	3840	3800	3840	3800	3840	3800

Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 2.28: Impact of learning and location of cookie choice (full table)

	Dependent variable: Cookie choice $\in \{0, 1\}$					
	Privacy concern × Learning		Privacy concern × Location		Privacy concern × (Learning + Location)	
	(1)	(2)	(3)	(4)	(5)	(6)
Exclusive	0.891**	0.948*	0.585	0.624	0.931**	0.963*
	(0.409)	(0.513)	(0.417)	(0.517)	(0.442)	(0.540)
Considerate	-0.065	-0.264	0.279	0.074	0.181	-0.027
	(0.397)	(0.373)	(0.408)	(0.387)	(0.429)	(0.407)
Exclusive × Considerate	-0.900	-1.121**	-1.537***	-1.748***	-1.428**	-1.644***
	(0.568)	(0.526)	(0.580)	(0.542)	(0.613)	(0.574)
Second half	-0.142	-0.143			-0.144	-0.144
	(0.154)	(0.154)			(0.154)	(0.154)
Exclusive × Second half	-0.626***	-0.621***			-0.625***	-0.621***
	(0.229)	(0.228)			(0.229)	(0.229)
Considerate × Second half	0.187	0.187			0.194	0.193
	(0.215)	(0.214)			(0.215)	(0.215)
Exclusive × Considerate	-0.265	-0.282			-0.278	-0.293
× Second half	(0.313)	(0.313)			(0.314)	(0.314)
Mid			0.094	0.102	0.094	0.102
			(0.194)	(0.195)	(0.195)	(0.195)
Far			-0.041	-0.043	-0.042	-0.045
			(0.188)	(0.188)	(0.188)	(0.188)
Exclusive × Mid			-0.128	-0.125	-0.125	-0.121
			(0.280)	(0.280)	(0.282)	(0.282)
Exclusive × Far			-0.001	0.007	-0.000	0.007
			(0.278)	(0.278)	(0.281)	(0.280)
Concerned × Mid			-0.180	-0.182	-0.180	-0.181
			(0.272)	(0.271)	(0.272)	(0.272)
Considerate × Far			-0.561**	-0.549**	-0.561**	-0.548**
			(0.264)	(0.263)	(0.264)	(0.263)
Exclusive × Considerate			0.514	0.513	0.513	0.509
× Mid			(0.386)	(0.386)	(0.390)	(0.389)
Exclusive × Considerate			1.074***	1.063***	1.090***	1.078***
× Far			(0.383)	(0.383)	(0.386)	(0.386)
Market	No	Yes	No	Yes	No	Yes
Demographics	No	Yes	No	Yes	No	Yes
Iterative thinking	No	Yes	No	Yes	No	Yes
Observations	3840	3800	3840	3800	3840	3800

Standard errors in parentheses
* $p < .10$, ** $p < .05$, *** $p < .01$

Table 2.29: Interaction between privacy concern, learning and location

Chapter 3

Surge Demand in On-demand Platform

Based on Li (2020*b*).

3.1 Introduction

On-demand economy, sometimes also known as gig or sharing economy, has been growing rapidly around the world. It uses advanced algorithm via digital platforms to connect consumers' demands with the immediate provisions of the goods and services, which breaks the traditional rule of full-time working contract and offers the chance for self-employed or part-time workers to join the market in a flexible way. According to the recent report,[1] the size of the on-demand economy has reached USD 110 billion in 2019, compared to USD 92.8 billion in 2018. Participation in the on-demand economy has nearly tripled since 2016. In the field of transportation, Uber, Lyft and Blabla Car have formed a well-structured network of ride-sharing platform. Considering the accommodation service, Airbnb becomes the top choice for more and more families. Additionally, delivery service such as Postmates and Instacart or household tasks like TaskRabbit and Handy are all popular among customers (Einav et al., 2016).

Different from traditional economy, workers have much more flexibility in on-demand market and are able to choose their own working time. Chen et al. (2019) empirically show that Uber drivers are much better off with such real-time flexibility, by obtaining more than twice the surplus they would in a traditional environment. Meanwhile, in order to balance the uncertain demand with flexible supply, on-demand platform uses dynamic

[1] Please see https://rockresearch.com/on-demand-economy-research/ (accessed on December 18, 2020).

pricing to make the matching. For instance, Uber and Lyft usually raise their service rates in some extreme conditions such as rainy working days or weekend's midnights, so as to solve the problem of undersupply under such circumstance. These platforms take advantage of the flexibility provided in the on-demand service, and surge pricing is the most commonly used strategy to balance the demand and supply.

Based on the common sense, surge pricing should be applied when demand surges in a certain area and as a consequence more service providers are willing to go to this market. However, the empirical observations deviate from our expectation. Chen et al. (2015) analyze the data from Uber and show that although the surge pricing leads to a significant reduction in demand, the relevant effect on supply is ambiguous. Some service providers choose to move out of the area where the price surges. This argument is confirmed by the relevant report.[2] They also point out that several Uber drivers complain about no ride request when the surge pricing is applied. Thus, opposite to our common sense, not many service providers are attracted by the surge pricing and they are more willing to stay in their original area.

In order to explain these counter-intuitive observations, Guda and Subramanian (2019) introduce a model representing this situation and propose a new terminology "strategic surge pricing", which means that under certain circumstance the platform also raises the price when there are excess supply. The authors provide an innovative way to look at this problem, however, without considering the new entrants in their model. New entry to the market is an important factor considering the on-demand platform, since both the flexibility of choosing working time and choosing working place are key characteristics. Therefore in this chapter, we extend the existing model by taking the new entrants into consideration, focusing on an on-demand market where there are not only the fixed incumbents but also the flexible entrants. Our analysis mainly looks at the case in which the total supply are enough for the surge demand and would like to answer the following research questions: can the sufficient new entrants solve the problem of undersupply in the market with surge demand? How does the on-demand platform strategically react? And how does the welfare change accordingly?

Our study shows that even with sufficient supply, not all the new entrants enter the market with surge demand and thereby undersupply may still exist. This happens only when the new entry occurs at the first period and the aggregate supply does not exceed a threshold, which explains the aforementioned empirical observations. In order to solve the undersupply, we point out that the platform has the incentive to postpone the entry

[2] Please see https://www.sfgate.com/business/article/Report-says-Uber-surge-pricing-has-a-twist-some-6597012.php (accessed on December 18, 2020).

of the new entrants to the period when the demand surges. Since the new entrants increase the competition in the market, they impose the negative externalities on all the other service providers within the same area. Postponing the entry to the later period reduces the overall negative effects and therefore makes more workers enter the market later. We demonstrate that the platform gets more profits from the late entry, which also matches with social planner's interest. However, on the other hand, new entrants themselves prefer to enter earlier. Thus, we propose several methods to encourage late entry, including limiting the quota, using the bonus and assigning the matching priority to the existing incumbents. We also show that although the strategic surge pricing proposed by Guda and Subramanian (2019) can solve the problem of undersupply, it reduces the platform's profit as well. This partly explains those seemingly contradictory evidence from different papers.

3.2 Related Literature

This chapter contributes to the growing literature on on-demand platform. Einav et al. (2016) provide a good overview about on-demand market by explicitly pointing out the advantages of lowering entry costs for sellers and allowing individuals and small suppliers to compete with traditional providers of goods or service. They mix individual suppliers with the existing large firms, in order to identify how these small and flexible suppliers affect the existing firms and how the efficiency changes accordingly. However, they don't mention the application of surge pricing. Additionally, we differ from them by focusing on the individual sellers.

Several papers study the application of surge pricing in taxi industry, however, most of which analyze the surge pricing within one market. In other words, there is no possibility for service providers to move in-between, which limits the study on surge pricing. Banerjee et al. (2015) propose a threshold pricing mechanism based on the number of taxis or the relevant demands. When this number goes above or below a threshold, the prices will be different. They make a comparison between static pricing and threshold dynamic pricing with a queueing approach on the single market. Cachon et al. (2017) analyze a sequential game with self-scheduling capacity, which is also considered by this chapter. Self-control on the capacity is the main characteristic brought by the on-demand platform. Their paper focuses on the comparisons between different types of contracts by controlling the flexibility of prices and wages, and prove that the contract with dynamic prices and wages maximizes the platform's profit and surge pricing generally makes all the stakeholders better off. Ming et al. (2019) reach the same conclusion by analyzing the empirical data from a Chinese platform Didi. Their studies back up our focus on the flexibility of capacity, which are also emphasized by Castillo et al. (2017), Taylor (2018) and Gurvich et al. (2019). On the other hand, Zha et al. (2018) and (Bimpikis et al., 2019) try to understand the surge pricing from algorithm designer's perspective and propose different matching algorithms such as Cobb-Douglas matching to optimize the related mechanism, which are also commonly applied in the field of engineering and operations research. To the contrary, we treat the mechanism of surge pricing as given, but concentrate on the questions about why and when to apply the surge pricing. We want to analyze under which circumstance the platform applies the surge pricing and how the service providers react accordingly.

Another stream is related to the studies on spatial economics, where the movement between markets are taken into consideration. Buchholz (2018) constructs a dynamic model of spatial search and matching between consumers and providers, demonstrating the inefficiencies caused by search frictions and misallocation. Contrarily, Bimpikis et al. (2019) look at the situation where the service providers can freely move between different places, without incurring any costs. We differ from both studies by considering the costs of movement, which is a key factor regarding the trade-off between staying in the current market and moving to another area. Özkan and Ward (2020) discuss the dynamic matching in disjoint areas and propose an algorithm based on a continuous linear program. Guda and Subramanian (2019) is the paper closet to our study, building a model with two locations and two periods. Service providers can move between the markets, which stresses the flexibility of movement with the on-demand platform. We extend their model by adding new entrants in the market. We focus on how new entrants react with surge demand, mainly exploring the case in which the total supply is sufficient and checking whether enough new entrants could solve the problem of undersupply in the market with surge demand. A similar question was also discussed in (Guda and Subramanian, 2019), however, without considering the entry of new providers. We believe that new entrants play an important role concerning the on-demand service, since the flexibility of entering the market represents the biggest improvement with the on-demand platform. Thus, we contribute to this issue by checking when and where the new entrants enter the market and how the on-demand platform strategically reacts. We also explore whether the problem of undersupply still exists under such circumstance and how the welfare changes accordingly.

This work also contributes to the literature on two-sided markets (Rochet and Tirole, 2003, Armstrong, 2006*a*, Rochet and Tirole, 2006), where elasticities and cross-market externalities play an important role. In this chapter, we focus on the price changes caused by the demand and supply, not by the elasticity. Therefore, we just consider the cross-market externalities. However, main results will hold even if we take the elasticity-related issues into account.

Several empirical analyses also show some interesting facts that support our study. Chen et al. (2015) find that surge pricing is noisy, making a strong and negative impact on consumers' demands but a weak and positive effect on the supply. Deviating from the common sense, they also show that more service providers move out of the area when the price surges. It seems that surge pricing does not incentivize the providers in the way as designed. These facts constitute the starting point for this chapter. On the other hand, Chen and Sheldon (2015) find the evidence to support surge pricing which significantly leads to the higher supply of rides and increases the efficiency of the on-demand market. Farber (2015) demonstrates that service providers positively respond to both anticipated and unanticipated surge pricing. Both these findings don't contradict but complement our analysis, since our study figures out the special conditions, under which the counter-intuitive actions, such as the facts mentioned by Chen et al. (2015), take place. This chapter aims to understand such behavior and provide the relevant suggestions.

3.3 Model

We consider a setup following Guda and Subramanian (2019) and add new entrants into account. There are three types of players: consumers, independent service providers and on-demand platform. They are active in two adjacent markets, A and B, operating in two periods, t_1 and t_2.

In each period, there are certain initial demands and supplies in each market. New entrants decide whether to enter the platform. If the entry is determined, they also need to choose which market to enter. Otherwise, they can wait until the next period.[3] After observing the total demand and supply, the platform sets different prices in each market. Given the market price, service providers decide whether to stay and serve the current market, or move to the adjacent market. Only the service providers and consumers who are in the same market can be matched. Serving the customers or moving to another market lasts for one period.

[3] In our setting with two periods, new entrants have to choose between entering the platform at t_1 and at t_2.

Consumers

There is a continuum of consumers in each market, who require one unit of the service in each period. Their reservation price for the on-demand service is uniformly distributed over $[1, 2]$. In the first period, the number (mass) of consumers requiring service is the same in both markets, which is denoted as a. In the second period, one of the markets experiences a surge demand a_H and another faces the same demand a,[4] $a_H > a$. The probability of having a surge demand is equal for both markets. Let p_{ij} represent the price set by the platform in zone $i \in \{A, B\}$ at period $j \in \{1, 2\}$. Given p_{ij}, the number of consumers requiring the service in market i at period j is $D_{ij} = a_{ij}(2 - p_{ij})$, where $a_{ij} \in \{a, a_H\}$.

Providers

There is a pool of registered providers in each market. On a given day, some of the providers are available and start the service from the market they registered.[5] In order to model this randomization, we assume that the initial supply (N_A, N_B) is either (n_H, n_L) or (n_L, n_H), $n_H \geq n_L$ and both are equally likely. Additionally, there are totally n new entrants who would like to enter the platform. We can treat them as either those who have registered in the pool but initially did not plan to work on the given day or those who have never provided the service on the platform. However, due to some reasons such as the expectation of a high revenue in a rainy day, they decide to go back to or join in the platform. This represents the nature of flexibility, which is the key characteristic of the on-demand market. New entrants are so flexible that they can enter any market in any period. Nevertheless, after the entry, they will face the same condition as the incumbents.

In the beginning of each period, independent service providers observe the market price p_{ij} and decide whether to stay and serve the current market, or move to the adjacent market.[6] If the provider chooses to stay, he can serve the customers in the same market and the whole process takes one period. Otherwise, he could move to another market at the switching cost C_s, during which he will also spend one period and is not allowed to provide the service. Each independent provider incurs the fixed operating cost which is

[4]For instance, the area under some extreme weather experiences a surge demand.

[5]In reality, for example, the Uber drivers start to provide their service from the area where they are living. It also applies to the case in which the providers begin to work after finishing a specific assignment, such as sending their children to the school.

[6]Please note that the market price p_{ij} is set after the new entrants' decisions, therefore the new entrants in principle also need to decide whether to stay or not. However, they will never choose to move due to the extra switching cost that they have to pay.

normalized to zero. All the service providers maximize their expected profits in the two periods.

Platform

The platform uses the technological algorithm to match the consumers and service providers, in order to maximize its expected profit. In each period, after observing the aggregate demand and supply, the on-demand platform sets prices p_{ij} in different markets and matches the consumers with the providers in the same market. Platform gets a fixed portion $\lambda \in (0,1)$ from each transaction, which means that with the market price p_{ij}, the revenue a provider can keep is $(1-\lambda)p_{ij}$ and the profit shared to the platform is λp_{ij}.

If there exists oversupply or overdemand in a certain market, the randomized matching will be applied. For instance, if there are more available service providers in the market, the platform will randomly assign some providers to serve the consumers and the remaining will keep idle in this period. To the contrary, if the aggregate demand is more than the aggregate supply, not all the consumers will receive the service. Platform randomly selects consumers to match with the limited service providers.

Timing

It is a two-period game, in which discounting is not taken into account. In the beginning of the first period, nature decides initial supply (N_A, N_B) at t_1 and the market with surge demand at t_2. In the first period t_1, consumers who require the on-demand service join the platform, and new entrants decide whether and which market to enter. After observing the aggregate demand and supply, the platform sets p_{i1} in each market. Given the market price, independent service providers decide whether to stay or move to the adjacent market, meanwhile consumers determine whether to request the service. Based on these decisions, the platform matches consumers who require the service with the providers deciding to stay in the same market. At the end of the first period, providers who decide to stay complete the service, and those who choose to move reach the adjacent market.

In the second period, one of the markets experiences the surge demand by having more consumers join the platform. New entrants decide which market to enter if they haven't done it before. Given the total demand, the platform sets p_{i2} for different markets. Then potential consumers decide whether to require the service, and in the end, platform makes the matching between consumers and service providers.

Assumptions

Additionally, we make the following assumptions throughout the analysis:

$$(i) \quad a_H > n_H \geq n_L > a$$

$$(ii) \quad n_H + n_L + n \geq a_H + a$$

Assumption (i) states that without any new entrants and movements between the markets, there exists undersupply in the market that experiences the surge demand. No matter what the condition in that market is, the initial supply itself cannot satisfy the increased demand. To the contrary, Assumption (ii) guarantees that in the second period, the aggregate supply is sufficient to meet all the requests if they can be properly allocated in different markets. Overall, these two assumptions show that if there were a benevolent social planner who would like to solve the problem of undersupply in the market with surge demand, it would eventually work, by involving enough new entrants entering that area or letting sufficient incumbents move to that market. In our model, however, no such a benevolent social planner exists and all the players are incentivized by their own expected payoffs. Therefore, we aim to find out whether enough supply can solve such a problem and how the different players behave under such circumstance.

Moreover, without loss of generality, we assume that it is the market A that experiences the surge demand in the second period. Due to the symmetry, it is totally identical if the demand surges in market B. In order to simplify our analysis, in the following sections we concentrate on the case where the surge demand occurs in A.

3.4 Equilibrium Analysis

In this section, we analyze the equilibrium of this two-period game. Firstly, we discuss the benchmark case where the new entrants and incumbents cannot freely move. We check two scenarios: one with no permission of the new entry and movement between the markets, and another with such a permission fully controlled by the platform, which represents the first-best revenue solution. Next, we check the cases when all the players have the freedom to decide their own actions. Our analysis focuses on the case where each player knows which market will experience the surge demand in the second period. This is based on the fact that many service providers are so experienced that know the market condition well and meanwhile the on-demand platform will also send them the updates about the demand and supply correctly, especially concerning its reputation.[7] Under such circumstance new entrants know all the information they need, representing the best scenario for them. It is interesting to verify whether new entrants will eventually go to the area with surge demand and solve the problem of undersupply.

3.4.1 Benchmark

As a benchmark, we check two cases in which the new entrants and incumbents cannot make the decision by themselves. In the first case, neither the new entry nor the movement between the markets is allowed, which provides us a basic picture on how the surge pricing works under initial market condition. Then we look at the case where the platform decides the new entry and movement, in order to find out the optimal situation from the platform's perspective.

In the case without permission of the new entry and movement between the markets, two markets are completely independent. The initial supply and current demand in each market jointly determine the price. In the first period, no matter how the initial supply distributes, both markets have more supply than demand given that $n_H \geq n_L > a$. Thus, corresponding to the result of Guda and Subramanian (2019), the platform sets the price p_{i1} to maximize its profit in market i:

$$p_{i1} = \arg\max_{p_{i1} \geq 1} \ \lambda D_{i1} p_{i1} = \arg\max_{p_{i1} \geq 1} \ \lambda a(2 - p_{i1}) p_{i1} = 1.$$

The equation above shows that absent the capacity constraint the platform sets the price at 1. We define $\bar{p} = 1$ as the regular price, meaning that the platform always chooses the price of 1 if there are enough supply to serve the consumers.[8] In the second period, market B is the same as before and thereby the regular price $p_{B2} = \bar{p} = 1$ is still imposed. To the contrary, market A experiences the surge demand and the supply is not enough to satisfy all the requests. Thus, surge pricing is implemented, which is determined by

[7] Actually we can easily figure out that the platform has no incentive to not truthfully report the market with surge demand.

[8] This result is based on the assumption of linear demand function, in which the price elasticity of demand is not affected by the increase of demand in the second period. This means that if the total supply is enough the platform will set the regular price to maximize its profit, even in the market with surge demand. With this assumption we exclude the possibility that the surge demand makes the demand less elastic, leading to the higher price even with sufficient supply. Throughout the analysis, therefore we can focus on the price changes caused by the demand and supply, not by the elasticity itself. This is what we are interested in. However, using a more general demand function will not change our main results.

$$p_{A2} = \arg\max_{p_{A2}\geq 1} \ \lambda p_{A2}\min\{N_A, a_H(2-p_{A2})\} = 2 - \frac{N_A}{a_H}.^{9}$$

Overall, when the supply is sufficient, the market price is equal to $\bar{p}$. When the demand increases and exceeds the supply, such as the market A in the second period, the platform raises the price, which is commonly known as surge pricing.

Next, we look at the case when the platform fully controls the new entry and movement between the markets. As we have seen in the previous analysis, if there are enough providers in the market, setting the price at $\bar{p}$ maximizes the platform's profit. Given that the platform's profit is proportional to the market revenue, $\bar{p}$ also maximizes the total revenue. In this model with the assumption that $n_H + n_L + n \geq a_H + a$, even in the second period, the total supply is enough if the new entry is allowed. Thus, the platform should allow the new entry, and one of its optimal strategies could be that the platform let all the n new entrants enter market A at t_2 and also let $a_H - n - N_A$ service providers move from B to A at t_1.[10] With such a strategy, there will be a_H providers in market A and $N_A + N_B + n - a_H$ providers in market B at t_2. Since $N_A + N_B + n - a_H \geq a$, there are enough supplies in both markets, therefore the regular price will be applied and no surge pricing occurs in the second period. It is also straightforward to see that all the demands in the first period can be met. So this is the optimal situation for the platform, since the regular price is set at each market in two periods, that is, $p_{ij} = 1, i \in \{A, B\}, j \in \{1, 2\}$. Moreover, with the regular price, all the consumers are willing to request the service and finally all the demands will be satisfied. Thus, consumer surplus is also maximized in this case.

3.4.2 Enough New Entrants

In the following analysis, we discuss the cases where each player can freely choose his own action. New entrants decide when and where to enter the market, and incumbents choose whether to stay in the current place or move to the adjacent market. Regarding the demand uncertainty, we focus on the case that everyone knows which market will experience the surge demand. This matches with our common sense that the experienced service providers can correctly predict the market condition. Additionally, the platform also sends the updates about the current situation to all the providers. The platform has no reason to not report truthfully, since its profit is proportional to the market revenue.

[9] We can easily show that when p_{A2} exceeds the regular price $\bar{p}$, platform's profit decreases in the price. In optimality, demand is equal to supply, that is $N_A = a_H(2 - p_{A2})$.

[10] Note that if $a_H - n - N_A < 0$, it means that $-(a_H - n - N_A)$ service providers move from A to B at t_1. In principle there are various feasible strategies, and we just show one of the possibilities here.

The concern of reputation also acts as a repeated check for such information. Thus, it is reasonable to assume that all the players know exactly which market will experience the surge demand in the second period.

In order to keep the analysis interesting, we make a further assumption on the new entrants. We call that there are *enough new entrants* in the market if $n_L + n \geq a_H$. This restriction tells us that even with the low initial supply n_L in the zone experiencing the surge demand, if all the new entrants enter this market there will be no undersupply. As we can see, there are two types of movement in the markets, new entrants' entry and incumbents' movement between different markets. Enough new entrants ensure that without any incumbents' movement the new entry itself could solve the problem of undersupply if they would be properly located to the market with surge demand. Therefore, we are interested in whether the enough new entrants can solve such a problem, if they freely choose when and where to enter the market.

Throughout the following analysis, we focus on the subgame perfect equilibrium in pure strategies, which means that each new entrant uses a pure strategy to decide his entry and the incumbents also follow the pure strategy regarding whether to stay or move to another market. In equilibrium, each player can correctly anticipate all the others' actions in two periods, including new entrants' entry, incumbents' movement and platform's pricing strategies. We check two scenarios: when the new entrants enter at t_1 or enter at t_2. We also show that it will never be optimal that some of the new entrants enter at t_1 and some enter at t_2. In the end, we make the comparison between these scenarios, to figure out different players' optimal strategies.

3.4.2.1 New Entry at t_2

Let r_{ij} represent the expected revenue of a provider serving in market i at t_j. For an incumbent who starts to work from market A and decides to stay, he can get the total profits $r_{A1} + r_{A2}$ in two periods. If he decides to move to market B at t_1, his expected profit will be $r_{B2} - C_s$. We denote $\mu_i \in [0, 1]$ as the proportion of service providers in market i who decide to move at t_1. The number of the service providers moving from market A to B is $\mu_A N_A$. Similarly, $\mu_B N_B$ represents those who decide to move from market B to A. The expected revenue r_{ij} depends on the price and the market condition of demand and supply. We have $r_{ij} = (1-\lambda)p_{ij}min\{\frac{D_{ij}}{S_{ij}}, 1\}$, where S_{ij} is the aggregate supply who can actually serve the consumers in market i at t_j, and $min\{\frac{D_{ij}}{S_{ij}}, 1\}$ stands for the probability of serving the market i for each provider.

When the new entry happens at t_2, new entrants need to decide which market to go.

For a new entrant, he enters market A if $r_{A2} > r_{B2}$ and enters market B when $r_{A2} < r_{B2}$. On the other hand, incumbents have to determine whether to stay or move to another market at t_1. A switcher moves from B to A if $r_{B1} + r_{B2} < r_{A2} - C_s$ and moves from A to B when $r_{A1} + r_{A2} < r_{B2} - C_s$.

Initially without any movement between the markets, due to the undersupply in A, r_{A2} is larger than r_{B2} and new entrants prefer to go to market A.[11] With more new entrants' entering, r_{A2} will be driven down. The probability of serving goes below one when the aggregate supply exceeds the aggregate demand, that is, $S_{A2} > D_{A2}$. In the end, two scenarios are possible: one is that after all the new entrants' entering A, r_{A2} is still greater than or equal to r_{B2} and therefore all new entrants go to A; another is that with the new entry, r_{A2} decreases until it is equivalent to r_{B2}. After that the new entrants are split into two markets, however, as a steady state, $r_{A2} = r_{B2}$ always holds. Thus, we can conclude that the equilibrium condition of the new entrants is $r_{A2} \geq r_{B2}$.

Now we take the incumbents into account, to figure out the equilibrium.[12] Lemma 3.1 firstly characterizes the switching behavior.

Lemma 3.1. *When new entrants enter at t_2, μ_A and μ_B cannot be both positive. Only $\mu_B > 0$ is possible.*

Proof. See Appendix. □

Lemma 3.1 shows that no one is willing to move from A to B, since it is the common knowledge that market A will experience the surge demand at t_2. In order to maximize the expected profit, service providers may move to A or stay in their original place, but never move to B. Actually those switchers impose the negative externalities on the existing providers in the market they move to, and to the contrary, exert the positive externalities on those in the area they move from. Switchers increase the competition of the market they move to, and as a consequence, either lower the probability of serving

[11]Since there exists undersupply in A, $r_{A2} = (1-\lambda)p_{A2} > 1-\lambda$, with $p_{A2} > 1$. To the contrary, supply is sufficient in market B, therefore $r_{B2} = (1-\lambda)p_{B2}\frac{D_{B2}}{S_{B2}} < 1-\lambda$, due to the fact that $D_{B2} < S_{B2}$ and $p_{B2} = 1$.

[12]Please note that in order to find out the equilibrium, we first consider the new entrants and then the incumbents in the analysis. Assume towards the contradiction that we let the incumbents make the decision first. Given that they anticipate all the following actions correctly, they move from B to A iff $r_{B1} + r_{B2} < r_{A2} - C_s$. Both r_{B1} and r_{B2} increases in the number of the switchers and r_{A2} decreases in the number of switchers, therefore the equilibrium condition is $r_{B1} + r_{B2} = r_{A2} - C_s$. However, when the new entrants choose where to enter at t_2, r_{A2} is strictly larger than r_{B2} in the aforementioned equilibrium condition. New entrants will enter market A, which reduces r_{A2}. Then it becomes that $r_{B1} + r_{B2} > r_{A2} - C_s$ and for those who have moved from B to A, it is irrational to act in such a way. Overall, given that all the players are rational, we should consider the new entrants' actions first, and then check the incumbents' relevant strategies.

the consumers in that area if there were enough providers before their movement, or lower the market price if the supply was not sufficient. Thus, in either way, the existing providers in that market become worse off. Similarly, the switchers' moving out will benefit the providers in that market due to the less competition.

If it is the first scenario in which all the new entrants go to market A, in the equilibrium $r_{A2} \geq r_{B2}$ should hold. Considering the existing providers, if $r_{B1} + r_{B2} \geq r_{A2} - C_s$ there is no switcher moving between the markets; otherwise, some incumbents move from B to A and as an equilibrium $r_{B1} + r_{B2} = r_{A2} - C_s$. On the other hand, when the second scenario happens that not all the new entrants go to market A, the equilibrium condition is $r_{A2} = r_{B2}$, which means that $r_{B1} + r_{B2} > r_{A2} - C_s$. Therefore, no one moves between the markets in this scenario. Combining the equilibria in these two scenarios, we have the following result:

Lemma 3.2. *When the new entry happens at t_2, in the equilibrium there will be no undersupply in the market with surge demand.*

Proof. See Appendix. □

Lemma 3.2 shows that if enough new entrants enter the market at t_2, there will be no undersupply in market A. The reason is that with the surge demand at t_2, market A is more attractive than B and enough new entrants ensure the supply in this area. Corollary 3.1 summarizes the details about the equilibria:

Corollary 3.1. *When $N_B < \frac{N_A+n}{a_H}a$, not all the new entrants go to market A and no switcher moves between the markets. When $\frac{N_A+n}{a_H}a \leq N_B \leq \frac{2a(N_A+n)(1-\lambda)}{a_H(1-\lambda)-C_s(N_A+n)}$, all the new entrants enter market A and there is still no switcher moving in-between. When $N_B > \frac{2a(N_A+n)(1-\lambda)}{a_H(1-\lambda)-C_s(N_A+n)}$, all the new entrants go to market A and some switchers also move from B to A.*

Proof. See Appendix. □

Corollary 3.1 is intuitive, showing that when the existing providers in market B are not many, some new entrants go to B, although market A experiences the surge demand. This is because similar to the switchers, new entrants also impose the negative externalities on the market that they enter. Therefore, some new entrants will choose to enter a market without the surge demand, when the competition is not fierce. Under such circumstance, the equilibrium condition is $r_{A2} = r_{B2}$ and no one moves between the markets. On the other hand, when the number of the existing providers in market B exceeds a threshold such as $\frac{N_A+n}{a_H}a$ in this case, all the new entrants will avoid the competition in this area and go to the market with the surge demand. However, there is still no switcher moving in-between, since the benefit from the less competition cannot cover the loss of the profit at t_1 plus the switching cost. Only when the number of the existing providers in market B is large enough, that is $\frac{2a(N_A+n)(1-\lambda)}{a_H(1-\lambda)-C_s(N_A+n)}$ in this scenario, some incumbents esacpe from B to get more profits, even though they anticipate that all the new entrants will also go to market A.

Considering the welfare, new entry at t_2 benefits the consumers, since all of them can be served under such circumstance. Meanwhile, the platform sets the regular price on each market, maximizing not only its profit but the producer surplus as well. Proposition 3.1 summarizes the case when new entrants enter at t_2.

Proposition 3.1. *If new entrants enter at t_2, the problem of undersupply in the market with surge demand can be solved. First-best revenue solution can be attained, and both consumer surplus and platform's profit are maximized.*

3.4.2.2 New Entry at t_1

If the new entry happens at t_1, the situation becomes complicated because in principle both market A and B can be attractive to the new entrants. Thus, from new entrant's perspective, they need to compare the values of $r_{A1} + r_{A2}$ with $r_{B1} + r_{B2}$, to determine which market to enter. Before analyzing new entrants' strategic choices, we firstly characterize the existing providers' switching behavior.

Lemma 3.3. *When new entrants enter at t_1, μ_A and μ_B cannot be both positive. Only $\mu_B > 0$ is possible.*

Proof. See Appendix. □

Similar to the previous case, Lemma 3.3 shows that considering the incumbents, only the movement from B to A is possible, even though the new entry takes place at t_1. Unlike incumbents, new entrants have more freedom. They prefer to enter A if $r_{A1} + r_{A2} > r_{B1} + r_{B2}$ and enter B if $r_{A1} + r_{A2} < r_{B1} + r_{B2}$. In the following analysis, we check three possible scenarios separately from the new entrant's perspective: all enter market A, all enter market B, some go to market A and some go to market B.

We start with the case where all the new entrants enter market B. As an equilibrium condition we have that $r_{A1} + r_{A2} \leq r_{B1} + r_{B2}$. In this equilibrium, since $r_{B1} + r_{B2} > r_{A2} - C_s$, no switcher moves from B to A. Together with the fact that all the new entrants go to market B, there will be undersupply in market A in the second period.

Thus, surge pricing will be applied in A at t_2 and $r_{A2} > r_{B2}$. Moreover, under such circumstance, $r_{A1} = (1-\lambda)\frac{n}{N_A}$, which is greater than $r_{B1} = (1-\lambda)\frac{n}{N_B+n}$ by the assumption $N_B + n \geq a_H > N_A$. Therefore, $r_{A1} + r_{A2} \leq r_{B1} + r_{B2}$ cannot be true, and it is impossible for all the new entrants to enter market B. The reason is that the new entry to market B drives down both r_{B1} and r_{B2}. With the assumption of enough new entrants, if all of them enter market B, r_{B1} and r_{B2} will be lower than r_{A1}, which is definitely less than r_{A2} due to the problem of undersupply.

Lemma 3.4. *Only when $(N_A, N_B) = (n_H, n_L)$ and $n_H > n_L$, new entrants may prefer to go to market B. However, if this happens, not all of them enter market B and eventually as an equilibrium condition, $r_{A1} + r_{A2} = r_{B1} + r_{B2}$. There will be no switcher moving between the markets under such circumstance.*

Proof. See Appendix. □

Lemma 3.4 shows that although not all the new entrants go to B, under certain circumstance market B can also be their preference. This is only possible when the initial supply is low in market B. To the contrary, if market A has the lower intial supply and experiences the surge demand in the second period, market B will not be the new entrant's preference.

Next, if all the new entrants go to market A, $r_{A1} + r_{A2} \geq r_{B1} + r_{B2}$ should hold as an equilibrium condition. With the assumption of enough new entrant, we have that $(1-\lambda)\frac{a}{N_A+n} + (1-\lambda)\frac{a_H}{N_A+n} \geq 2(1-\lambda)\frac{a}{N_B}$. This means that $\frac{a+a_H}{N_A+n} \geq 2\frac{a}{N_B}$, which is equivalent to $N_B \geq \frac{2a(N_A+n)}{a+a_H}$. When there are too many existing service providers in market B, all the new entrants will choose to go to A, since the total probabilities of serving the consumers in the two periods are larger in market A. Similar to Corollary 3.1, we can show that there is a threshold, above which some incumbents move from B to A and under which no switcher moves between the markets. In either case, there is sufficient total supply in market A at t_2. Lemma 3.5 summarizes the details in this scenario:

Lemma 3.5. *When $N_B \geq \frac{2a(N_A+n)}{a+a_H}$, all the new entrants go to market A. If $\frac{2a(N_A+n)}{a+a_H} \leq N_B \leq \frac{2a(1-\lambda)(N_A+n)}{(1-\lambda)a_H - C_s(N_A+n)}$, there is no switcher moving between the markets. If $N_B > \frac{2a(1-\lambda)(N_A+n)}{(1-\lambda)a_H - C_s(N_A+n)}$, some existing providers move from B to A.*

Proof. See Appendix. □

Finally, we check the scenario in which some new entrants go to market A and some go to market B. No matter what the new entrant's preference is initially, he should be indifferent between going to market A and B in the equilibrium. Thus, the equilibrium condition is $r_{A1} + r_{A2} = r_{B1} + r_{B2}$, and there is no switcher moving from B to A since $r_{B1} + r_{B2} > r_{A2} - C_s$ under such circumstance. Considering the fact that not all the new entrants enter market A in this scenario, it may be possible that the total supply is not sufficient and thereby the undersupply still exists in market A at t_2. Lemma 3.6 describes such a possibility:

Lemma 3.6. *When the total supply in the second period ($n_H + n_L + n$) is less than $a_H + \frac{2a_H}{a+a_H}a$, there will be undersupply in market A.*

Proof. See Appendix. □

Lemma 3.6 shows that enough new entrants cannot ensure the sufficient supply in the market with surge demand. This happens when the total supply exceeds the total demand within a certain range. Although there are enough new entrants, due to the negative externalities they bring to the area that they enter, not all of them go to the same market. Thus, when the excess supply is not too many, new entrants cannot solve the problem of undersupply.

In our setup, everyone knows the market that will experience the surge demand in the second period. The existing service providers are able to freely move between the markets and the new entrants can enter any market according to their preferences. Together with the assumption of enough new entrants, it seems that either the new entrants or the existing service providers will fill the gap between the demand and supply. However, we show that the problem of undersupply still exists under certain circumstance.

If we compare this result with Lemma 3.2, we can see that such a problem would be solved when the new entry happens at t_2. As we mentioned before, both the new entrants and switchers impose the negative externalities on the market that they intend to go. If the new entry takes place in the second period, such negative externalities just exist for one period. For the new entrants, the benefit from the surge demand exceeds the negative externalities brought by all the other new entrants in this period. Therefore, they are willing to enter the market with surge demand until the problem of undersupply has been solved. The assumption of enough new entrants ensures this result. However, when the new entry occurs at t_1, such negative externalities would be amplified, since they will influence each provider in both periods. In other words, the benefit from the surge demand is weakened and the high profit in the second period may not attract enough new entrants. This is true when the total supply is not too large, meaning that the competition in two markets is not fierce and some new entrants are willing to enter the market without surge demand. On the other hand, if the aggregate supply exceeds the threshold, the competition will be intense and the new entrants should take advantage of the surge demand in the second period. In this condition, there will be no undersupply in the market with surge demand. The following proposition summarizes the case when the new entry occurs at t_1.

Proposition 3.2. *Even with enough aggregate supply, not all the new entrants enter the market with surge demand, and thereby undersupply may exist. This only takes place when the new entrants enter at t_1 and the total supply at t_2 is less than $a_H + \frac{2a_H}{a+a_H}a$.*

3.4.2.3 Welfare

In the previous cases, we have shown that if the new entry happens at t_2, the aggregate supply is sufficient in the market with surge demand. To the contrary, if the new entry takes place at t_1, the problem of undersupply may exist. In this section, we make the comparisons regarding new entrant's payoff, platform's profit as well as social surplus between these two cases, to fully understand the differences in incentives among different players.

First, we look at the new entrants who can freely choose when and where to enter the market. Previously we have discussed the cases in which all the new entrants enter at t_1 or t_2. Before making the comparison between these two cases, we need the following result as a supplement.

Lemma 3.7. *It is not an optimal strategy that some new entrants enter at t_1 and some enter at t_2.*

Proof. See Appendix. □

Lemma 3.7 helps us rule out the possibility that different new entrants enter in different periods and focus on the aforementioned cases where the entry takes place at t_1 or t_2. Now we can compare the new entrant's payoff to decide when they prefer to enter the market.

Lemma 3.8. *From new entrant's perspective, they prefer to enter the market in the first period.*

Proof. See Appendix. □

Lemma 3.7 and 3.8 show that if the new entrants have the freedom to decide when to enter, they will enter the market in the first period. However, we have shown that if the new entry happens at t_1, the problem of undersupply may still exist. To the contrary, the new entry at t_2 could solve such a problem. Therefore, we need to check the platform's incentive as well as the social planner's strategy, to verify whether they are consistent with the new entrants' preferences.

Considering the platform, its profit is a fixed portion to the market revenue. Therefore, from its perspective, the platform wants to make as many matchings between consumers and service providers as possible. When all the demands have been satisfied, the regular price will be set in the market and the platform reaches its optimal situation, no matter how the matchings are eventually made among the players. Thus, it is straightforward to see that the platform prefers the new entry at t_2, which contradicts to the new entrants' incentives. Meanwhile, the regular price ensures that each consumer is willing to request the service and all such demands will be met in the end. So consumer surplus is also maximized in this case. Since both the platform's profit and consumer surplus are maximized under the regular price, as well as the market revenue, we only need to check the total costs to find out the social optimal condition. Lemma 3.9 provides the relevant result.

Lemma 3.9. *Social welfare is greater when the new entry happens at t_2, which coincides with the platform's incentive.*

Proof. See Appendix. □

The analysis shows that new entrants have the different incentives with the platform and social planner. The new entrants prefer to enter the market in the first period, in order to obtain more profits. To the contrary, both the platform and social planner are in favor of the late entry, which could solve the problem of undersupply in the second period and maximize the platform's profit as well as the social welfare.

New entrants bring the negative externalities to the market they enter. When the new entry happens at t_2, such negative externalities just last for one period. However, they will influence the both periods when the new entry takes place at t_1. Therefore, from social planner's perspective, the optimal condition is to minimize the effects caused by such negative externalities. Meanwhile, new entry at t_2 can also solve the problems of undersupply brought by the surge demand. That's why the social planner prefers the new entry at t_2.

On the other hand, from new entrant's perspective, they do not internalize the negative externalities that they impose on the other service providers in the same market. When entering the market, they will have a certain positive probability to serve the consumers, which in return increases their expected profits. The benefit from early entry exceeds the loss of profit from the intense competition, and therefore new entrants prefer to enter the market in the first period.

3.5 Discussion

Based on the aforementioned analysis, we have seen that new entrants prefer to enter the on-demand market in the beginning, in order to obtain more expected profits. To the contrary, the platform has the incentive to postpone their entry to the period when the demand surges, in order to solve the problem of undersupply. Therefore, in this section, we want to point out the possible methods for the platform to solve such a problem, together with a brief discussion regarding opportunity costs.

Strategic Surge Pricing

Strategic surge pricing, initially proposed by Guda and Subramanian (2019), means that under certain circumstance the platform can raise the price in the market with the excess supply, that is market B in this case, to solve the problem of undersupply. Recall that when the entry happens in the first period, the new entrants enter market A iff $r_{A1} + r_{A2} > r_{B1} + r_{B2}$. In order to make more entrants go to market A, the platform could increase p_{B1}. By doing so, the expected profit of entering market B is reduced and therefore more entrants are willing to go to another market. Proposition 3.3 describes such a strategy.

Proposition 3.3. *In order to solve the problem of undersupply when demand surges, the platform can use the surge pricing on the market with oversupply in the first period. He can raise p_{B1} above a threshold, that is $1 + \sqrt{2 - \frac{(n_H+n_L+n-a_H)(a+a_H)}{aa_H}}$ in this case, in order to prevent the entry of the new entrants to the market without surge demand.*

Proof. See Appendix. □

This explains the empirical observations that several service providers complain about no ride requests when the price is surging. From the proposition we can see that the aim of the surge pricing is not to satisfy the increased demands, but lower the probability of serving in that area. Thus, the expected profits decrease, incumbents have less customers to serve and new entrants go to another market where demand surges. This is the reason that not many providers are attracted by the surge pricing and they incline to stay in where they are.

Although the strategic surge pricing seems to be attractive concerning the problem of undersupply, we still have to check whether it can indeed benefit the platform. By increasing p_{B1}, the platform needs to make a trade-off between increasing the profit in market A at t_2 and losing the payoff in market B at t_1. With our setup, the following result shows the profit change from platform's perspective.

Corollary 3.2. *Although strategic surge pricing can solve the problem of undersupply, it also lowers down the platform's profits.*

Proof. See Appendix. □

Unlike Guda and Subramanian (2019) where the authors show that strategic surge pricing can be profitable under certain circumstance, Corollary 3.2 demonstrates that it is not the case when the new entrants are taken into account. Although this method can make more new entrants go to market A and satisfy the surge demands, the costs in market B are higher. Therefore, theoretically speaking, the platform has no incentive to follow such a strategy. The difference between these results is mainly caused by the different purposes of implementing the strategy. In Guda and Subramanian (2019), they would like to incentivize more switchers move from the market with oversupply to the market with surge demand. However, our purpose is to prevent the new entrants from entering the market without surge demand.

On the other hand, if we take the factors such as reputation and confidence into consideration, it may be possible for the platform to use strategic surge pricing. The reason is that if the platform always couldn't satisfy the surge demands, consumers will lose the confidence in this on-demand market. Therefore, under certain condition, the platform may implement the strategic surge pricing, which explains those seemingly contradictory empirical evidence from different papers (Chen et al., 2015, Farber, 2015, Chen and Sheldon, 2015). However, since it exceeds the scope of this chapter, we will leave the further discussion for the future research.

Policy Implication

In the previous part we have shown that strategic surge pricing can solve the undersupply but does not coincide with the platform's interest. Thus, we would like to point out other feasible methods to fix this problem.

As demonstrated above, new entrants choose to enter the market in the first period and therefore the undersupply may exist in the market with surge demand in the second period. From platform's perspective, the direct way to solve this problem is to encourage the late entry or postpone the entry to the second period. Following this logic, the platform can limit the quota of entry for the first period or offer the priority to the incumbents, such as assigning the matching priority to the existing service providers. Under such circumstance, the new entrants are not allowed to enter or afraid of being idle during the first period with the limited demands. Thus, some of them will choose the late entry in the second period, which can solve the undersupply at t_2. As a practice, the delay of informing or limiting the success rate of matching among new entrants can be carried out. For the platform, this is the least costly method to deal with the problem of undersupply.

On the other hand, even if the new entry happens in the first period, the plaform can still incentivize more entrants to go to the market with the surge demand. One method is to offer the bonus to those who enter the market with the surge demand. In practice, we find the similar method to support this idea. For instance, Uber launches a bonus program called Consecutive Trips, in which the service provider will get a bonus if he completes a certain number of rides in a row during a specific time period and the first ride must start in a specified zone.[13] In our case, it means that the platform encourages the providers to serve market A in order to satisfy the surge demand in the second period.[14]

Opportunity Costs

In the analysis we ignore the opportunity cost that each service provider incurs. Considering the on-demand platform, it affects the trade-off between joining in the platform or dealing with certain personal arrangement. Regarding the new entrants, this trade-off is crucial, since they may not have the plan to join the platform initially. If the opportunity cost is higher than the expected profit, provider will not serve the consumer in that period and postpone his entry. This could in turn solve the problem of the undersupply in the second period, which coincides with our interest.

In order to identify the effect of opportunity cost on the new entrant's behavior, we assume that for each service provider, his opportunity cost is identical in two periods, but heterogeneous among all the providers. We firstly check the difference of the expected profit between two periods if the new entry happens at t_1. Lemma 3.10 describes the relevant result.

[13] Please see: https://help.uber.com/driving-and-delivering/article/how-does-the-consecutive-trips-promotion-work?nodeId=de983305-076a-40cf-aaf4-7b23f50a0007 (accessed on December 18, 2020).

[14] In principle, when to start serving market A is not important. The platform just needs to ensure the sufficient supply when the demand surges. In Lemma 3.1 and 3.3 we have proven that no service providers will move from A to B, that is, $\mu_A = 0$.

Lemma 3.10. *When entering at t_1, no matter which market the new entrant goes to, the expected profit at t_1 cannot be larger than that at t_2.*

Proof. See Appendix. □

When the opportunity cost is large, especially larger than the expected profit at t_1 but less than the profit at t_2, new entrants will choose the late entry. Considering the heterogeneous opportunity costs for different service providers, the tendency would be that more providers enter at t_2, which can at least partly solve the problem of undersupply when demand surges at t_2.

3.6 Conclusion

This chapter provides a simple analysis about the on-demand platform. We use a two-period two-market model to identify the effect of surge demand on the decisions about new entrant's entry and incumbent's movement. We show that even with sufficient supply, not all the new entrants will enter the market with surge demand and therefore the undersupply still exists under certain circumstance. The platform has the incentive to postpone the entry of new entrants to the period when the demand surges, in order to solve the undersupply and obtain more profits. This coincides with the social planner's interest, but conflicts with the new entrant's preference. Thus, we propose several ways to deal with this problem, including the strategic surge pricing. Differing from the findings by Guda and Subramanian (2019), we demonstrate that although the strategic surge pricing can solve the problem of undersupply, it lowers down the platform's profit as well.

Throughout the analysis we focus on the common knowledge, meaning that the information about the initial supply and the market that will experience surge demand is commonly known by all the players. As we mentioned before, this is reasonable concerning the service provider's experience and platform's reputation. However, it would be interesting to extend this model to an asymmetric case, in which only the on-demand platform has such information. Under this condition, it is easy to see that the platform has no incentive to not truthfully report the market with surge demand, since its profit is proportional to the market revenue. To the contrary, it might deserve a further study on how to signal the initial supply from the platform's perspective, in order to maximize the expected profits.

We also would like to point out some issues which may draw attention in the future research. The price discrimination is not considered in this model. However, as mentioned by Bimpikis et al. (2019), the spatial pricing in ride-sharing networks should be a trend that deserves the attention. Moreover, introducing multiple periods or more states in the model may lead to different results, which would be interesting to investigate. Additionally, if we replace the linear demand by a general form, especially in which the price elasticity changes with the surge demand, it should provide further insights on surge pricing in the on-demand platform. However, our conjecture is that the main results should still hold with a more general demand function.

3.7 Appendix

Proof of Lemma 3.1

If $\mu_A > 0$ it means that some providers move from A to B. Therefore, in the first period, a service provider is willing to move if and only if $r_{A1} + r_{A2} < r_{B2} - C_s$. Similarly, $\mu_B > 0$ means that $r_{B1} + r_{B2} < r_{A2} - C_s$. Since these two inequalities cannot hold simultaneously, μ_A and μ_B cannot be both positive. Assume towards a contradiction that $\mu_A > 0$, so $r_{A1} + r_{A2} < r_{B2} - C_s$ should hold. Since no one moves from B to A under such circumstance, together with the surge demand in A at t_2, there exists undersupply in A without new entrants. Therefore, $r_{A2} > r_{B2}$ and new entrants should enter market A. For the new entrants, as we have shown, the equilibrium condition is $r_{A2} \geq r_{B2}$. This is a contradiction to the initial assumption that $r_{A1} + r_{A2} < r_{B2} - C_s$. Thus, only $\mu_B > 0$ is possible.

Proof of Lemma 3.2

As analyzed in the main text, there are two scenarios concerning the equilibrium. If all the new entrants enter market A, there will be no undersupply due to the assumption of enough new entrants. If not all the new entrants go to market A, the equilibrium condition is $r_{A2} = r_{B2}$ and no one moves from B to A. Since there are sufficient supply in market B, the probability of serving the consumers is below one and the regular price is set. Thus, $r_{B2} < 1 - \lambda$, and the same for r_{A2}. We can show that only when the undersupply does not exist, the revenue can be lower than $1 - \lambda$. Assume towards a contradiction that the undersupply still exists in market A, then the probability of serving the consumers should be one and the price is above one. So the revenue of the individual provider is larger than $1 - \lambda$, which contradicts to the initial assumption. Thus, concerning $r_{A2} = r_{B2} < 1 - \lambda$, there should be no undersupply in market A. Combining these two scenarios we can conclude that there will be no undersupply in the market with surge demand.

Proof of Corollary 3.1

From Lemma 3.2 we have got that there is no undersupply in market A. Thus, the price in each market at each period will be 1. If all the new entrants go to market A, $(1-\lambda)\frac{a_H}{N_A+n} \geq (1-\lambda)\frac{a}{N_B}$ should hold, which is equivalent to $N_B \geq \frac{N_A+n}{a_H}a$. Under such circumstance, when $r_{B1}+r_{B2} \geq r_{A2}-C_s$ there is no switcher moving between the markets, which we have $(1-\lambda)\frac{2a}{N_B} \geq (1-\lambda)\frac{a_H}{N_A+n} - C_s$ and solve that $N_B \leq \frac{2a(N_A+n)(1-\lambda)}{a_H(1-\lambda)-C_s(N_A+n)}$. To the contrary, when $r_{B1}+r_{B2} < r_{A2}-C_s$, some switchers move from B to A. We have $(1-\lambda)\frac{2a}{N_B} < (1-\lambda)\frac{a_H}{N_A+n} - C_s$, and derive that $N_B > \frac{2a(N_A+n)(1-\lambda)}{a_H(1-\lambda)-C_s(N_A+n)}$. On the other hand, if not all the new entrants enter A, $N_B < \frac{N_A+n}{a_H}a$ which can be directly derived from $(1-\lambda)\frac{a_H}{N_A+n} < (1-\lambda)\frac{a}{N_B}$.

Proof of Lemma 3.3

The method of proof is similar to Lemma 3.1. $\mu_A > 0$ means that $r_{A1}+r_{A2} < r_{B2}-C_s$, and $\mu_B > 0$ is equivalent to $r_{B1}+r_{B2} < r_{A2}-C_s$. However, these two inequalities cannot hold simultaneously. So μ_A and μ_B cannot be both positive. Assume towards a contradiction that $\mu_A > 0$, we have $r_{A1}+r_{A2} < r_{B2}-C_s$. Since no existing provider moves from B to A under such circumstance, without new entrants, there would be undersupply in A and oversupply in B, resulting in $r_{A2} > r_{B2}$. The only possibility to make $r_{A1}+r_{A2} < r_{B2}-C_s$ hold is that in the beginning of t_1 new entrants prefer to enter market A. Therefore, we should have $r_{A1}+r_{A2} \geq r_{B1}+r_{B2}$ as the equilibrium condition, together with $r_{A1}+r_{A2} \leq r_{B2}-C_s$ by assumption. However, these two inequalities cannot be true simultaneously, which proves that only $\mu_B > 0$ is possible.

Proof of Lemma 3.4

For a new entrant, he prefers to enter market B iff $r_{B1}+r_{B2} > r_{A1}+r_{A2}$ when $\mu_B = 0$. This is equivalent to $2(1-\lambda)\frac{a}{N_B} > (1-\lambda)\frac{a}{N_A} + (1-\lambda)(2-\frac{N_A}{a_H})$, where we can get that $\frac{2a}{N_B} - \frac{a}{N_A} + \frac{N_A}{a_H} > 2$. Since $\frac{N_A}{a_H} < 1$, we have $\frac{2a}{N_B} - \frac{a}{N_A} > 1$. This can be true only when $(N_A, N_B) = (n_H, n_L)$ and $n_H > n_L$.

Proof of Lemma 3.5

We have shown that when $N_B \geq \frac{2a(N_A+n)}{a+a_H}$, all the new entrants go to market A. Under such circumstance, if $r_{B1}+r_{B2} \geq r_{A2}-C_s$, there will be no switcher moving between the markets. Thus, we have that $(1-\lambda)\frac{2a}{N_B} \geq (1-\lambda)\frac{a_H}{N_A+n} - C_s$, which is equivalent to $N_B \leq \frac{2a(1-\lambda)(N_A+n)}{(1-\lambda)a_H - C_s(N_A+n)}$. On the other hand, if $r_{B1}+r_{B2} < r_{A2}-C_s$, some existing providers will move from B to A. We can easily get that N_B should be large enough to satisfy this condition, that is, $N_B > \frac{2a(1-\lambda)(N_A+n)}{(1-\lambda)a_H - C_s(N_A+n)}$.

Proof of Lemma 3.6

Assume that n_A new entrants enter A, and n_B new entrants enter B, where $n_A + n_B = n$. Since there is no switcher moving between the markets, in the second period the total supply at market A is $N_A + n_A$ and the total supply at market B is $N_B + n_B$. As we have analyzed, the equilibrium condition is $r_{B1} + r_{B2} = r_{A1} + r_{A2}$, which can be written as

$$(1-\lambda)\frac{2a}{N_B + n_B} = (1-\lambda)\frac{a}{N_A + n_A} + (1-\lambda)p_{A2}\min\{1, \frac{a_H}{N_A + n_A}\}.$$

If the total supply in market A is not sufficient, it means that $N_A + n_A < a_H$. Under such circumstance, $p_{A2} = 2 - \frac{N_A + n_A}{a_H} > 1$ and the probability of serving the consumers at A is one. Therefore, the equilibrium condition above is equivalent to

$$(1-\lambda)\frac{2a}{N_B + n_B} = (1-\lambda)\frac{a}{N_A + n_A} + (1-\lambda)(2 - \frac{N_A + n_A}{a_H}).$$

Together with the constraint $N_A + n_A < a_H$, we can get that $\frac{2a}{N_B + n_B} > \frac{a}{a_H} + 1$. Additionally, considering that the problem of undersupply still exists in market A, the aggregate supply in B should be more than $n_L + n_H + n - a_H$. In other words, $N_B + n_B > n_L + n_H + n - a_H$. Combining these two inequalities, we can derive that $n_H + n_L + n < a_H + \frac{2a_H}{a + a_H}a$.

Proof of Lemma 3.7

Assume towards the contradiction that some new entrants enter at t_1 and some enter at t_2. Then they are indifferent between these two choices and should not enter the same market in two periods. For instance, if some new entrants enter market A at t_1, all those who choose to enter at t_2 will go to market B, and vice versa. The new entrants who enter at t_2 receive less profit than those who enter at t_1 in the same market, since the expected profit in the first period is strictly positive. Thus, there are just two possibilities here: some entering A at t_1 and some entering B at t_2, or some entering B at t_1 and some entering A at t_2. In the first scenario, $r_{A1} + r_{A2} = r_{B2}$ as the equilibrium condition. However, we have $r_{A1} + r_{A2} < r_{B1} + r_{B2}$ under such circumstance, which means that those who enter at t_1 will deviate to go to market B. Similarly, in the second scenario, $r_{B1} + r_{B2} = r_{A2}$ should hold in the equilibrium, where we have $r_{B1} + r_{B2} < r_{A1} + r_{A2}$

and therefore the new entrants should prefer to go to market A at t_1. These deviations show the contradiction to the initial assumption, which complete the proof.

Proof of Lemma 3.8

In order to prove that the new entrants prefer to enter in the first period, our method is to firstly list out all the scenarios if they enter at t_2 and then show that they have the incentive to deviate in each scenario.

In the previous analysis we have shown that if the new entry happens at t_2 there will be two scenarios: all the new entrants go to A and not all the new entrants go to A. In the first scenario where all the new entrants enter market A, the equilibrium condition is $r_{A2} \geq r_{B2}$. However, if they deviate to enter the same market at t_1, this would not affect market B but make themselves better off. The reason is as follows: for the switchers who may move from B to A, this deviation has no influence, since they compare the values of $r_{B1} + r_{B2}$ with $r_{A2} - C_s$ and the deviation brings no effect on these two values. Basically, entering at t_1 changes the number of the providers in market A at t_1 and therefore lowers r_{A1}. However, r_{A2} will keep the same, since there are still the same number of providers in market A at t_2. The same logic could be applied to market B, because the deviation has no influence on the number of providers in market B in both periods. Therefore, both r_{B1} and r_{B2} will not change, and the deviation makes no influence on the incumbents in market B. On the other hand, the existing service providers in market A become worse off due to the decrease of r_{A1}. But they have no incentive to move to B, since $r_{A1} + r_{A2} > r_{B2} - C_s$ with $r_{A2} \geq r_{B2}$ still holding in the equilibrium. To the contrary, the new entrants who deviate to enter the same market at t_1 will be strictly better off because they can get the extra revenue r_{A1} while keeping r_{A2} the same as before. Overall, with the deviation, new entrants become better off, the incumbents in market A are worse off and the incumbents in market B are indifferent. Although it may not be the best strategy if the new entrants enter at t_1, this is a feasible and profitable deviation for the new entrants. Thus, the first scenario is not optimal from the new entrant's perspective.

In the second scenario where not all the new entrants enter market A at t_2, we have shown that the equilibrium condition is $r_{A2} = r_{B2}$ and no switcher moves between the markets. If the new entrants deviate to enter the same market at t_1, there still would be no switcher moving in-between, since $r_{B1} + r_{B2} > r_{A2} - C_s$ with $r_{A2} = r_{B2}$ holding in the equilibrium. Thus, the number of the providers in market B at t_2 does not change. However, since some new entrants go to B at t_1 with the deviation, r_{B1} would be driven down and the incumbents in market B become worse off. Considering the existing providers in market A, they would also be worse off with the deviation, because more providers at t_1 lowers r_{A1}. However, the same as the previous analysis, they have no incentive to deviate to market B, since $r_{A1} + r_{A2} > r_{B2} - C_s$ with $r_{A2} = r_{B2}$. For the new entrants, they are strictly better off with the deviation due to the extra revenue they could receive at t_1. Please note that this would not be the optimal situation, since $r_{A1} + r_{A2}$ may not be equivalent to $r_{B1} + r_{B2}$ under such circumstance and some new entrants should make the further deviation. Anyways, whatever the optimal strategy is, the new entrants would benefit from the deviation. Combining these two scenarios, we can show that the new entrants prefer to enter in the first period.

Proof of Lemma 3.9

In order to prove Lemma 3.9, we list out all the scenarios when the new entry happens at t_1 and make the comparison with the cases if the new entrants would enter at t_2. We show that in each scenario, the social surplus would not be lower if the new entrants could enter at t_2 and under certain circumstance it would become strictly larger. The same as the analysis in the main text, two cases are discussed separately when the new entry takes place at t_1: all the new entrants go to market A, or not all the new entrants go to market A.

We first look at the case in which not all the new entrants go to market A. The equilibrium condition is $r_{A1}+r_{A2} = r_{B1}+r_{B2}$ and no switcher moves between the markets. In this scenario if the new entrants enter at t_2, their initial preferences would be market A. However, with more and more new entrants entering market A, two possibilities may occur. The first one is that all the new entrants go to market A with $r_{A2} \geq r_{B2}$ in the equilibrium. We also have $r_{B1} + r_{B2} > r_{A2} - C_s$ here, since both r_{B1} and r_{B2} increase with fewer new entrants going to B and r_{A2} decreases with more new entrants going to A. Thus, there is no switcher moving from B to A in this situation. Another possibility is that some new entrants go to A and some go to B with $r_{A2} = r_{B2}$ in the equilibrium. Under such circumstance, we have $r_{B1} + r_{B2} > r_{A2} - C_s$ and therefore still no switcher moves in-between. Overall, if we compare all the possibilities between entering at t_1 and entering at t_2 in this case, we can conclude that there will be no switcher under such condition, so the total costs are the same. However, as shown in Lemma 3.6, there may exist the undersupply in market A if the new entry takes place at t_1. Thus, considering the social welfare, entering at t_2 is better than entering at t_1 in this case.

Next, we check the case where all the new entrants go to market A. From Lemma 3.5

we have verified that there are two scenarios in this case: no switcher and with switcher.

If there is no switcher moving between the markets, the equilibrium condition is $r_{B1}+r_{B2} \geq r_{A2}-C_s$ and $r_{B1}+r_{B2} \leq r_{A1}+r_{A2}$. In this scenario if the new entry happens at t_2, there are two possibilities. The first one is that all the new entrants enter market A with $r_{A2} \geq r_{B2}$ in the equilibrium. Under such circumstance, $r_{B1}+r_{B2} \geq r_{A2}-C_s$ still holds, since all the r_{B1}, r_{B2} and r_{A2} do not change compared with the case if entering at t_1 and therefore no switcher moves in-between. The second possibility is that not all the new entrants go to A with $r_{A2} = r_{B2}$ in the equilibrium, where we can easily see that $r_{B1}+r_{B2} > r_{A2}-C_s$ and there is no switcher. To sum it up, we show that no switcher moves from B to A if the new entry happens at t_2, which is the same as the condition when the new entry occurs at t_1. By making the comparison between the two cases, we can see that no matter when the new entry takes place, the problem of undersupply in market A at t_2 will be solved and therefore both the consumer surplus and market revenue will be the same. Thus, although the final equilibrium conditions may be different, the social surplus does not change with respect to the time of entry.

On the other hand, if there are some switchers moving from B to A, the equilibrium condition is $r_{B1}+r_{B2} = r_{A2}-C_s$ and $r_{B1}+r_{B2} < r_{A1}+r_{A2}$. If the new entrants enter at t_2 in this scenario, all would go to market A. This is because we have $r_{B1}+r_{B2} < r_{A2}-C_s$ if all the new entrants enter market A and no switcher moves from B to A, where we can easily get that $r_{B2} < r_{A2}$ even all the new entrants have entered A. Since all the new entrants go to market A no matter when the new entry happens, r_{B1} and r_{B2} are independent of the entry. Together with the fact that r_{A2} will not be affected, the number of switchers do not change in two cases. So we can identify that in both cases all the new entrants and switchers act in the same way, and the social surplus will also be the same.

Combining all the scenarios mentioned above, we can conclude that entering at t_2 makes the social welfare at least the same as entering at t_1 in all the scenarios and under certain conditions strictly better off. Therefore, social welfare is greater when the new entry happens at t_2.

Proof of Proposition 3.3

When the new entrants enter in the first period, they choose market A iff $r_{A1}+r_{A2} > r_{B1}+r_{B2}$. As analyzed in the main text, the platform can increase p_{B1} in order to reduce the profit of entering market B. Thus, more entrants would go to market A. As an equilibrium condition, we have $r_{A1}+r_{A2} = r_{B1}+r_{B2}$. We also need $N_A+n_A \geq a_H$, to ensure the sufficient supply in market A at t_2. Moreover, the higher p_{B1} is, the lower profit the platform could get and the more new entrants go to market A. Thus, in the equilibrium the inequality should be binding, that is, $N_A + n_A = a_H$. By plugging the revenues into the equilibrium condition $r_{A1} + r_{A2} = r_{B1} + r_{B2}$ we have

$$(1-\lambda)p_{B1}\frac{a(2-p_{B1})}{N_B+n_B} + (1-\lambda)\frac{a}{N_B+n_B} = (1-\lambda)\frac{a+a_H}{N_A+n_A}.$$

Please note that with the surge pricing p_{B1}, for the service provider the probability of serving in market B at t_1 is $\frac{a(2-p_{B1})}{N_B+n_B}$. Since $N_A + n_A = a_H$, inserting $N_B + n_B = n_H + n_L + n - a_H$ into the above equation, we derive

$$p_{B1}\frac{a(2-p_{B1})}{n_H+n_L+n-a_H} + \frac{a}{n_H+n_L+n-a_H} = \frac{a+a_H}{a_H}.$$

By solving this quadratic equation we can get that

$$p_{B1} = 1 + \sqrt{2 - \frac{(n_H+n_L+n-a_H)(a+a_H)}{aa_H}} \in (1,2).$$

Proof of Corollary 3.2

We focus on the case mentioned by Lemma 3.6, where the undersupply exists with enough new entrants. Without strategic surge pricing, the platform's profit is:

$$\pi = \lambda[p_{A1}\cdot a + p_{A2}\cdot a_H \cdot (N_A+n_A) + p_{B1}\cdot a + p_{B2}\cdot a] = \lambda[3a + p_{A2}(N_A+n_A)].$$

With strategic surge pricing, the platform's profit will be:

$$\pi' = \lambda[p'_{A1}\cdot a + p'_{A2}\cdot a_H + p'_{B1}\cdot a\cdot(2-p'_{B1}) + p'_{B2}\cdot a = \lambda[2a + a_H + p'_{B1}(2-p'_{B1})a].$$

Thus, we get $\pi - \pi' = \lambda[a + p_{A2}(N_A+n_A) - a_H - p'_{B1}(2-p'_{B1})a]$. Concerning the proof of Lemma 3.6, as the equilibrium condition we have $r_{B1}+r_{B2} = r_{A1}+r_{A2}$, which is equivalent to $\frac{2a}{N_B+n_B} = \frac{a}{N_A+n_A} + p_{A2}$. Therefore, $p_{A2} = \frac{2a}{N_B+n_B} - \frac{a}{N_A+n_A}$. On the other hand, from the proof of Proposition 3.3 we have got that $p'_{B1}(2-p'_{B1})a = \frac{a+a_H}{a_H}(n_H+n_L+n-a_H) - a$. Inserting these two results into the equation above, we have that

$$\begin{aligned}\pi - \pi' &= \lambda[a + p_{A2}(N_A+n_A) - a_H - p'_{B1}(2-p'_{B1})a] \\ &= \lambda[2a\frac{N_A+n_A+N_B+n_B}{N_B+n_B}] - (n_H+n_L+n)\frac{a+a_H}{a_H}.\end{aligned}$$

Since $N_A+n_A+N_B+n_B = n_H+n_L+n$, we get that $\pi-\pi' = \lambda(n_H+n_L+n)(\frac{2a}{N_B+n_B} - \frac{a+a_H}{a_H})$.

Moreover, based on the proof on Lemma 3.6, $\frac{2a}{N_B+n_B} > \frac{a+a_H}{a_H}$ is the prerequisite for the existence of the undersupply with enough new entrants. Thus, $\pi > \pi'$.

Proof of Lemma 3.10

In order to figure out the profit difference in two periods, we consider two cases separately when the new entry happens at t_1: all the new entrants go to market A, or not all the new entrants go to market A.

When all the new entrants go to market A, the equilibrium condition is $r_{A1} + r_{A2} \geq r_{B1} + r_{B2}$. As we have shown, there are two possibilities regarding the switchers. If there is no switcher moving between the markets, we have that $r_{A1} = (1-\lambda)\frac{a}{N_A+n} < r_{A2} = (1-\lambda)\frac{a_H}{N_A+n}$. If there are switchers moving in-between, we have that $r_{A1} = (1-\lambda)\frac{a}{N_A+n}$, $r_{A2} = (1-\lambda)\frac{a_H}{N_A+n+\tilde{n_B}}$ and $r_{B1} = r_{B2} = (1-\lambda)\frac{a}{N_B-\tilde{n_B}}$, where $\tilde{n_B}$ represents the number of providers moving from B to A at t_1. Since $N_A + n > N_B - \tilde{n_B}$ by the assumption of enough new entrants, we have that $r_{A1} < r_{B1} = r_{B2}$. Combining with the equilibrium condition $r_{A1} + r_{A2} \geq r_{B1} + r_{B2}$, we get that $r_{A1} < r_{B1} = r_{B2} < r_{A2}$. Therefore, we show that under this circumstance the expected profit at t_2 is larger than that at t_1.

When not all the new entrants go to market A, the equilibrium condition is $r_{A1}+r_{A2} = r_{B1} + r_{B2}$. There is no switcher moving between the markets since $r_{B1} + r_{B2} > r_{A2} - C_s$ in this case. Thus, for those who enter market B, $r_{B1} = r_{B2} = (1-\lambda)\frac{a}{N_B+n_B}$ where n_B stands for the number of new entrants entering market B. We also have $r_{A1} = (1-\lambda)\frac{a}{N_A+n_A} < r_{A2} = (1-\lambda)\frac{a_H}{N_A+n_A}$ where n_A represents the number of new entrants entering market A. Overall, we can conclude that when the new entry happens at t_1, no matter which market the new entrant enters, the expected profit at t_1 cannot be larger than that at t_2.

Bibliography

Ackley, G. (1942), 'Spatial competition in a discontinuous market', *The Quarterly Journal of Economics* **56**(2), 212–230.

Acquisti, A. and Varian, H. R. (2005), 'Conditioning prices on purchase history', *Marketing Science* **24**(3), 367–381.

Acquisti, A., John, L. K. and Loewenstein, G. (2013), 'What is privacy worth?', *The Journal of Legal Studies* **42**(2), 249–274.

Acquisti, A., Taylor, C. and Wagman, L. (2016), 'The economics of privacy', *Journal of Economic Literature* **54**(2), 442–92.

Agrawal, A., Catalini, C. and Goldfarb, A. (2014), 'Some simple economics of crowdfunding', *Innovation Policy and the Economy* **14**(1), 63–97.

Ahlers, G. K., Cumming, D., Günther, C. and Schweizer, D. (2015), 'Signaling in equity crowdfunding', *Entrepreneurship Theory and Practice* **39**(4), 955–980.

Ali, S. N., Lewis, G. and Vasserman, S. (2019), Voluntary disclosure and personalized pricing, Working paper, National Bureau of Economic Research.

Armstrong, M. (2006*a*), 'Competition in two-sided markets', *The RAND Journal of Economics* **37**(3), 668–691.

Armstrong, M. (2006*b*), Recent developments in the economics of price discrimination, Cambridge University Press.

Banerjee, S., Riquelme, C. and Johari, R. (2015), 'Pricing in ride-share platforms: A queueing-theoretic approach', *Available at SSRN 2568258* .

Barreda-Tarrazona, I., García-Gallego, A., Georgantzís, N., Andaluz-Funcia, J. and Gil-Sanz, A. (2011), 'An experiment on spatial competition with endogenous pricing', *International Journal of Industrial Organization* **29**(1), 74–83.

Baye, I. and Sapi, G. (2014), *Targeted pricing, consumer myopia and investment in customer-tracking technology*, number 131, DICE Discussion Paper.

Belleflamme, P. and Vergote, W. (2016), 'Monopoly price discrimination and privacy: The hidden cost of hiding', *Economics Letters* **149**, 141–144.

Belleflamme, P., Lambert, T. and Schwienbacher, A. (2014), 'Crowdfunding: Tapping the right crowd', *Journal of Business Venturing* **29**(5), 585–609.

Belleflamme, P., Omrani, N. and Peitz, M. (2015), 'The economics of crowdfunding platforms', *Information Economics and Policy* **33**, 11–28.

Beresford, A. R., Kübler, D. and Preibusch, S. (2012), 'Unwillingness to pay for privacy: A field experiment', *Economics Letters* **117**(1), 25–27.

Bimpikis, K., Candogan, O. and Saban, D. (2019), 'Spatial pricing in ride-sharing networks', *Operations Research* **67**(3), 744–769.

Block, J., Hornuf, L. and Moritz, A. (2018), 'Which updates during an equity crowdfunding campaign increase crowd participation?', *Small Business Economics* **50**(1), 3–27.

Braghieri, L. (2019), 'Targeted advertising and price discrimination in intermediated online markets', *Available at SSRN 3072692* .

Brokesova, Z., Deck, C. and Peliova, J. (2014), 'Experimenting with purchase history based price discrimination', *International Journal of Industrial Organization* **37**, 229–237.

Buchholz, N. (2018), Spatial equilibrium, search frictions and dynamic efficiency in the taxi industry, Working paper, Working Paper.

Cachon, G. P., Daniels, K. M. and Lobel, R. (2017), 'The role of surge pricing on a service platform with self-scheduling capacity', *Manufacturing & Service Operations Management* **19**(3), 368–384.

Camacho-Cuena, E., García-Gallego, A., Georgantzís, N. and Sabater-Grande, G. (2005), 'Buyer–seller interaction in experimental spatial markets', *Regional Science and Urban Economics* **35**(2), 89–108.

Carroni, E. et al. (2015), *Competitive behaviour-based price discrimination among asymmetric firms*, CERPE.

Casadesus-Masanell, R. and Hervas-Drane, A. (2015), 'Competing with privacy', *Management Science* **61**(1), 229–246.

Castillo, J. C., Knoepfle, D. and Weyl, G. (2017), Surge pricing solves the wild goose chase, *in* 'Proceedings of the 2017 ACM Conference on Economics and Computation', pp. 241–242.

Chang, J.-W. (2020), 'The economics of crowdfunding', *American Economic Journal: Microeconomics* **12**(2), 257–80.

Chemla, G. and Tinn, K. (2020), 'Learning through crowdfunding', *Management Science* **66**(5), 1783–1801.

Chen, L., Mislove, A. and Wilson, C. (2015), Peeking beneath the hood of uber, *in* 'Proceedings of the 2015 Internet Measurement Conference', pp. 495–508.

Chen, M. K. and Sheldon, M. (2015), 'Dynamic pricing in a labor market: Surge pricing and the supply of uber driver-partners', *University of California (Los Angeles) Working Paper URL http://citeseerx. ist. psu. edu/viewdoc/download* .

Chen, M. K., Rossi, P. E., Chevalier, J. A. and Oehlsen, E. (2019), 'The value of flexible work: Evidence from uber drivers', *Journal of Political Economy* **127**(6), 2735–2794.

Cho, I.-K. and Kreps, D. M. (1987), 'Signaling games and stable equilibria', *The Quarterly Journal of Economics* **102**(2), 179–221.

Colombo, S. (2016), 'Imperfect behavior-based price discrimination', *Journal of Economics & Management Strategy* **25**(3), 563–583.

Conitzer, V., Taylor, C. R. and Wagman, L. (2012), 'Hide and seek: Costly consumer privacy in a market with repeat purchases', *Marketing Science* **31**(2), 277–292.

Cumming, D. J., Hornuf, L., Karami, M. and Schweizer, D. (2020), 'Disentangling crowdfunding from fraudfunding', *Max Planck Institute for Innovation & Competition Research Paper* (16-09).

Cumming, D. J., Leboeuf, G. and Schwienbacher, A. (2020), 'Crowdfunding models: Keep-it-all vs. all-or-nothing', *Financial Management* **49**(2), 331–360.

Dinev, T. and Hart, P. (2006), 'An extended privacy calculus model for e-commerce transactions', *Information Systems Research* **17**(1), 61–80.

Dufwenberg, M., Sundaram, R. and Butler, D. J. (2010), 'Epiphany in the game of 21', *Journal of Economic Behavior & Organization* **75**(2), 132–143.

Einav, L., Farronato, C. and Levin, J. (2016), 'Peer-to-peer markets', *Annual Review of Economics* **8**, 615–635.

Ellman, M. and Hurkens, S. (2019), 'Optimal crowdfunding design', *Journal of Economic Theory* **184**, 104939.

Esteves, R.-B. (2014), 'Price discrimination with private and imperfect information', *The Scandinavian Journal of Economics* **116**(3), 766–796.

Esteves, R. B. et al. (2009), A survey on the economics of behaviour-based price discrimination, Working paper, NIPE-Universidade do Minho.

European Commission (2020), 'A European strategy for data'. https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf.

Farber, H. S. (2015), 'Why you can't find a taxi in the rain and other labor supply lessons from cab drivers', *The Quarterly Journal of Economics* **130**(4), 1975–2026.

Feri, F., Giannetti, C. and Jentzsch, N. (2016), 'Disclosure of personal information under risk of privacy shocks', *Journal of Economic Behavior & Organization* **123**, 138–148.

Fischbacher, U. (2007), 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics* **10**(2), 171–178.

Fudenberg, D. and Tirole, J. (2000), 'Customer poaching and brand switching', *RAND Journal of Economics* pp. 634–657.

Fudenberg, D. and Villas-Boas, J. M. (2006), 'Behavior-based price discrimination and customer recognition', *Handbook on Economics and Information Systems* **1**, 377–436.

Ghosh, A., Mahdian, M., McAfee, R. P. and Vassilvitskii, S. (2015), 'To match or not to match: Economics of cookie matching in online advertising', *ACM Transactions on Economics and Computation (TEAC)* **3**(2), 1–18.

Gneezy, U., Rustichini, A. and Vostroknutov, A. (2010), 'Experience and insight in the race game', *Journal of Economic Behavior & Organization* **75**(2), 144–155.

Greiner, B. (2015), 'Subject pool recruitment procedures: Organizing experiments with orsee', *Journal of the Economic Science Association* **1**(1), 114–125.

Guda, H. and Subramanian, U. (2019), 'Your uber is arriving: Managing on-demand workers through surge pricing, forecast communication, and worker incentives', *Management Science* **65**(5), 1995–2014.

Gurvich, I., Lariviere, M. and Moreno, A. (2019), Operations in the on-demand economy: Staffing services with self-scheduling capacity, *in* 'Sharing Economy', Springer, pp. 249–278.

Hakenes, H. and Schlegel, F. (2014), 'Exploiting the financial wisdom of the crowd–crowdfunding as a tool to aggregate vague information', *SSRN 2475025* .

Heiny, F., Li, T. and Tolksdorf, M. (2020), 'We value your privacy: Behavior-based pricing under endogenous privacy', *Available at SSRN 3508762* .

Hotelling, H. (1929), 'Stability in competition', *The Economic Journal* **39**(153), 41–57.

Kunz, M. M., Bretschneider, U., Erler, M. and Leimeister, J. M. (2017), 'An empirical investigation of signaling in reward-based crowdfunding', *Electronic Commerce Research* **17**(3), 425–461.

Li, T. (2020*a*), 'Crowdfunding with asymmetric information', *Working paper* .

Li, T. (2020*b*), 'Surge demand in on-demand platform', *Working paper* .

Liu, Q. and Serfes, K. (2004), 'Quality of information and oligopolistic price discrimination', *Journal of Economics & Management Strategy* **13**(4), 671–702.

Loertscher, S. and Marx, L. M. (2020), 'Digital monopolies: Privacy protection or price regulation?', *International Journal of Industrial Organization* p. 102623.

Mahmood, A. (2014), 'How do customer characteristics impact behavior-based price discrimination? An experimental investigation', *Journal of Strategic Marketing* **22**(6), 530–547.

Malhotra, N. K., Kim, S. S. and Agarwal, J. (2004), 'Internet users' information privacy concerns (iuipc): The construct, the scale, and a causal model', *Information Systems Research* **15**(4), 336–355.

Mikians, J., Gyarmati, L., Erramilli, V. and Laoutaris, N. (2012), Detecting price and search discrimination on the internet, *in* 'Proceedings of the 11th ACM workshop on hot topics in networks', pp. 79–84.

Mikians, J., Gyarmati, L., Erramilli, V. and Laoutaris, N. (2013), Crowd-assisted search for price discrimination in e-commerce: First results, *in* 'Proceedings of the ninth ACM Conference on Emerging Networking Experiments and Technologies', pp. 1–6.

Ming, L., Tunca, T. I., Xu, Y. and Zhu, W. (2019), 'An empirical analysis of market formation, pricing, and revenue sharing in ride-hailing services', *Pricing, and Revenue Sharing in Ride-Hailing Services (February 15, 2019)* .

Mollick, E. (2014), 'The dynamics of crowdfunding: An exploratory study', *Journal of Business Venturing* **29**(1), 1–16.

Mollick, E. R. and Kuppuswamy, V. (2014), 'After the campaign: Outcomes of crowdfunding', *UNC Kenan-Flagler Research Paper* (2376997).

Montes, R., Sand-Zantman, W. and Valletti, T. (2018), 'The value of personal information in online markets with endogenous privacy', *Management Science* .

Özkan, E. and Ward, A. R. (2020), 'Dynamic matching for real-time ride sharing', *Stochastic Systems* **10**(1), 29–70.

Parliament and Council of the European Union (2016), 'Regulation (eu) 2016/679'. `http://data.europa.eu/eli/reg/2016/679/oj` accessed May, 6 2020.

Preibusch, S., Kübler, D. and Beresford, A. R. (2013), 'Price versus privacy: An experiment into the competitive advantage of collecting less personal information', *Electronic Commerce Research* **13**(4), 423–455.

Rochet, J.-C. and Tirole, J. (2003), 'Platform competition in two-sided markets', *Journal of the European Economic Association* **1**(4), 990–1029.

Rochet, J.-C. and Tirole, J. (2006), 'Two-sided markets: A progress report', *The RAND Journal of Economics* **37**(3), 645–667.

Schudy, S. and Utikal, V. (2017), 'You must not know about me–on the willingness to share personal data', *Journal of Economic Behavior & Organization* **141**, 1–13.

Schwienbacher, A. (2018), 'Entrepreneurial risk-taking in crowdfunding campaigns', *Small Business Economics* **51**(4), 843–859.

Shilony, Y. (1977), 'Mixed pricing in oligopoly', *Journal of Economic Theory* **14**(2), 373–388.

Strausz, R. (2017), 'A theory of crowdfunding: A mechanism design approach with demand uncertainty and moral hazard', *American Economic Review* **107**(6), 1430–76.

Streitfeld, D. (2000), 'On the web price tags blur: What you pay could depend on who you are', *The Washington Post* (September 27).

Taylor, C. R. (2004), 'Consumer privacy and the market for customer information', *RAND Journal of Economics* pp. 631–650.

Taylor, T. A. (2018), 'On-demand service platforms', *Manufacturing & Service Operations Management* **20**(4), 704–720.

Tsai, J. Y., Egelman, S., Cranor, L. and Acquisti, A. (2011), 'The effect of online privacy information on purchasing behavior: An experimental study', *Information Systems Research* **22**(2), 254–268.

Tucker, C. E. (2012), 'The economics of advertising and privacy', *International Journal of Industrial Organization* **30**(3), 326–329.

Zha, L., Yin, Y. and Du, Y. (2018), 'Surge pricing and labor supply in the ride-sourcing market', *Transportation Research Part B: Methodological* **117**, 708–722.

Selbständigkeitserklärung

Ich bezeuge durch meine Unterschrift, dass meine Angaben über die bei der Abfassung meiner Dissertation benutzten Hilfsmittel, über die mir zuteil gewordene Hilfe sowie über frühere Begutachtungen meiner Dissertation in jeder Hinsicht der Wahrheit entsprechen.

Berlin, den 31. Dezember 2020

Tianchi Li